

**CARACTERIZACIÓN DE LA MORBILIDAD BASADA EN TÉCNICAS DE
ANÁLISIS DE DATOS PARA HOSPITALES NIVEL III EN EL MUNICIPIO DE
MEDELLÍN**

Modalidad: Exploratorio

**JUAN MARTÍN VÉLEZ
TOMÁS RESTREPO MONTOYA**

Trabajo de grado para optar al título de Ingeniero Industrial

Director: Jaime Alberto Sánchez

Profesor de Cátedra e Ingeniero Industrial



**UNIVERSIDAD EIA
INGENIERÍA INDUSTRIAL
ENVIGADO
2019**

AGRADECIMIENTOS

Queremos agradecer a Sebastián Bedoya por la ayuda brindada en la realización de este trabajo. Además, a María Camila Ochoa quién aportó ideas para guiar el estudio de series de tiempo.

CONTENIDO

	pág.
INTRODUCCIÓN.....	12
1. PRELIMINARES.....	13
1.1 Planteamiento del problema	13
1.2 Objetivos del proyecto	17
1.2.1 Objetivo General.....	17
1.2.2 Objetivos Específicos	17
1.3 Marco de referencia.....	17
2. METODOLOGÍA.....	30
3. PRESENTACIÓN Y DISCUSIÓN DE RESULTADOS.....	71
4. CONCLUSIONES Y CONSIDERACIONES FINALES	88
REFERENCIAS	90
ANEXO 1	91
ANEXO 2.....	97
ANEXO 3.....	101

LISTA DE TABLAS

	pág.
TABLA 1 Variables Cualitativas de los RIPS.....	20
TABLA 2 Variables Cualitativas de los RIPS.....	21
TABLA 3 Ejemplo datos RIPS extraídos de la base de datos.....	32
TABLA 4 Ejemplo datos filtrados y modificados.....	34
TABLA 5 Clasificación grupos por género y edad vital.....	36
TABLA 6 Demanda mensual para el ID procedimiento 9073 con finalidad Terapéutico para el Grupo 8 con diagnóstico O800 en el año 2017.....	61
TABLA 7 Demanda mensual para el ID procedimiento 12211 con finalidad Diagnóstico para el diagnóstico Z139 para la población general.....	61
TABLA 8 Distribución empírica discreta de la severidad.....	65
TABLA 9 Días de estancia e histograma.....	74
TABLA 10 Resultados morbilidad general.....	76
TABLA 11 Segmentación por grupos.....	77
TABLA 12 Frecuencia enfermedad Z139 por municipio de procedencia del paciente.....	78
TABLA 13 Clasificación ABC de los procedimientos de Z139.....	79
TABLA 14 : Frecuencia enfermedad O800 por municipio de procedencia del paciente...	80
TABLA 15: Clasificación ABC de los procedimientos de O800.....	80

LISTA DE FIGURAS

	pág.
FIGURA 1 Secuencia en UiPath.....	23
FIGURA 2 Esquema de Flowchart UiPath.....	24
FIGURA 3 Ejemplo Series de Tiempo.....	26
FIGURA 4 Series de Tiempo Multiplicativa y Aditivas.....	27
FIGURA 5 Interpretación OpVar.....	28
FIGURA 6 Representación gráfica del LAD.....	29
FIGURA 7 Ejemplo campos de la base de datos SISPRO.....	30
FIGURA 8 Actividades 1, 2, 3 y 4 Clasificación morbilidad.....	37
FIGURA 9 For Each Row actividad 5.....	39
FIGURA 10 Condicionales pasos 5a, 5b y 5c.....	40
FIGURA 11 Actividades 6 y 7.....	41
FIGURA 12 Actividades 8 - 12.....	42
FIGURA 13 Etapa 2 Filtrar enfermedades.....	44
FIGURA 14 Actividad 1.....	45
FIGURA 15 Actividades a, b, c y d.....	46
FIGURA 16 Actividades e y f.....	48
FIGURA 17 Actividades g.i., g.ii. y g.iii.....	49
FIGURA 18 Actividad g.iv.....	49
FIGURA 19 Actividad g.v.....	51
FIGURA 20 Actividad g.vi.....	52
FIGURA 21 Actividad h.....	53

FIGURA 22 Actividad lectura Prueba_DatosRips.xlsx.....	54
FIGURA 23 Literales 2 y 3.....	55
FIGURA 24 Configuración de la actividad Filter Data Table.....	56
FIGURA 25 Literales 1 y 2.....	58
FIGURA 26 Literales 3 – 7.....	59
FIGURA 27 Estadística descriptiva demanda mensual (Frecuencia) ID 9073 procedimiento con finalidad terapéutica para el año 2017.....	62
FIGURA 28 Estadística descriptiva para la severidad del 2017 del procedimiento 9073 con finalidad terapéutica.....	62
FIGURA 29 Histograma para el comportamiento de la demanda del procedimiento 9073.....	63
FIGURA 30 Histograma de los costos para el procedimiento 9073 analizado.....	64
FIGURA 31 Pruebas de bondad de ajuste demanda procedimiento 9073 analizado.....	64
FIGURA 32 Prueba de bondad de ajuste Severidad 9073.....	65
FIGURA 33 Ajuste de LPA según Risk Utilizando AIC.....	66
FIGURA 34 Estadística descriptiva frecuencia procedimiento 12211 analizado.....	67
FIGURA 35 Histograma de frecuencia procedimiento 12211.....	67
FIGURA 36 Medida de ajuste de distribución geométrica.....	67
FIGURA 37 Histograma procedimiento 12211.....	68
FIGURA 38 Prueba de bondad de Ajuste para severidad.....	69
FIGURA 39 Distribución discreta empírica.....	69
FIGURA 40 Ajuste LPA según Risk Utilizando AIC para 12211.....	70
FIGURA 41 Porcentaje por género.....	71
FIGURA 42 Porcentaje por edad y género.....	71
FIGURA 43 Porcentaje por grupo etario.....	72

FIGURA 44 Tipo de usuario.....	72
FIGURA 45 Tipo de atención.....	73
FIGURA 46 Mortalidad.....	73
FIGURA 47 Personal que atiende.....	74
FIGURA 48 Gráfica series de tiempo totales por semana.....	81
FIGURA 49 Serie de tiempo de enfermedad Z139.....	82
FIGURA 50 Serie de tiempo de procedimiento 12211.....	82
FIGURA 51 Descomposición de tendencia de Z139 y 12211.....	83
FIGURA 52 Descomposición de estacionalidad de Z139 y 12211.....	83
FIGURA 53 Serie de tiempo de enfermedad O800.....	84
FIGURA 54 Serie de tiempo de procedimiento 9073.....	85
FIGURA 55 Descomposición de tendencia de O800 y 9073.....	85
FIGURA 56 Descomposición de estacionalidad de O800 y 9073.....	86

LISTA DE ANEXOS

	pág.
ANEXO 1 Código SQL.....	94
ANEXO 2 Código R.....	97
ANEXO 3 Datos Severidad.....	101

GLOSARIO

1. Morbilidad: Cantidad de personas que se enferman en un lugar y un período de tiempo determinados en relación con el total de la población.
2. RIPS: Registros Individuales de Prestación de Servicios de Salud; datos generados en cada visita de un paciente a una entidad de prestación de servicios de salud.

RESUMEN

La logística hospitalaria moderna se torna cada día más compleja debido al incremento en el número de pacientes que requieren de atención especializada y las cambiantes problemáticas que presentan. Esto representa un gran desafío para lograr entender la demanda a la cual se enfrentan los prestadores de servicios de salud. A estas dificultades se suma el interés del gobierno por obtener una mayor cobertura de salud y su necesidad por mantener los costos de los IPS cada día más bajos. Hasta el momento la logística hospitalaria en Colombia no ha recibido la atención que requiere y además no cuenta con herramientas que sean capaces de caracterizar de forma eficaz la naturaleza dinámica de su demanda. Este estudio, tratará de proponer una metodología para la caracterización efectiva de la demanda o patrones identificables de la demanda, en los modelos de logística hospitalaria para hospitales nivel III en la ciudad de Medellín.

Se efectuará el estudio mediante el trazo de series de tiempo, herramientas potentes que permiten identificar los aspectos claves de la demanda, utilizando los datos RIPS del Ministerio de Salud. La metodología utilizada es en gran parte de autoría propia, fundamentada con varias herramientas y técnicas aprendidas durante la carrera de Ingeniería Industrial en la EIA.

Se espera obtener una caracterización de la morbilidad actual de los hospitales nivel III en la ciudad para entender las enfermedades que causan la mayor congestión en los centros de prestación de servicios de salud. Posteriormente, se espera una identificación de los procedimientos más costosos para Hospitales Nivel III y finalmente una caracterización por series de tiempo en el uso de estos procedimientos y sus respectivas enfermedades.

Palabras Claves: Logística hospitalaria; hospital nivel III; series de tiempo; caracterización de la morbilidad

ABSTRACT

The modern health logistic chain becomes every day more and more complex due to the large increase in the number of patients which require specialized attention and due to ever changing problems. This represents a huge challenge in understanding the actual demand to which health services providers are faced. Adding to these difficulties, we add the interest from governments to achieve a broader healthcare coverage and the necessity to maintain, every day, lower costs. Up to this day, healthcare logistics has not received the attention it should and does not count with the tools needed to characterize effectively the dynamic nature of its demand. This study, will try to present a methodology to characterize demand or its identifiable patterns, effectively, on healthcare logistic models for level III hospitals in the city of Medellin.

The study will be performed through time series analysis, potent tools which enable the identification of key aspects of demand, using data from the Colombian Health Ministry digging into databases called RIPS. The proposed methodology is greatly from own propositions, with foundations on various tools and techniques learnt during the industrial engineering career at EIA University.

An actual morbidity characterization for level III hospitals in the city is expected as a result of the study, to better understand the diseases which cause mayor congestion on health service providers. Subsequently, a procedure identification for key procedures to deal with these diseases and finally a characterization by time series in the use of these procedures and its respective diseases are expected.

Keywords: Health logistic chain, level III hospital; time series; morbidity characterization

INTRODUCCIÓN

El análisis de datos mediante herramientas informáticas ha tomado importantísima relevancia en los últimos años debido a la ventaja de procesar muchos datos para obtener conclusiones y extraer información de forma mucho más precisa. Estas técnicas se están utilizando en la actualidad a través de todos los campos de la economía y programas sociales para incrementar la efectividad en la toma de decisiones. En el sector de la salud alrededor del mundo se utilizan comúnmente para aumentar la efectividad de los programas de prevención, desarrollar nuevos medicamentos y tratamientos y conocer a fondo las causas de mortalidad de las personas. La OMS “analiza los datos de sus 193 Estados Miembros y elabora estimaciones de carga de morbilidad y mortalidad que se publican en el informe Estadísticas Sanitarias Mundiales”, para indicar como cambian los riesgos para la salud en el mundo.

Estos estudios no han sido realizados en la misma medida en Colombia. El objetivo de este trabajo, por ende, busca definir una metodología utilizando datos ya existentes, para analizar la morbilidad de la ciudad de Medellín y analizar los procedimientos que se utilizan a la hora de tratar estas enfermedades.

Las diferentes etapas del trabajo mostrarán una clasificación de la morbilidad, seguido de una clasificación ABC de los procedimientos más utilizados para tratar las enfermedades en hospitales de nivel III y por último, el análisis en series de tiempo de ambas para encontrar patrones específicos que puedan llevar a una mejor planeación en los modelos de logística hospitalaria en la ciudad de Medellín.

1. PRELIMINARES

1.1 PLANTEAMIENTO DEL PROBLEMA

1.1.1. Formulación del Problema

Como mencionan (Chopra & Meindl, 2016), el éxito o fracaso de las organizaciones depende en gran parte del diseño y la gestión de la cadena de suministro, es decir una gestión logística exitosa. Lograr la gestión logística adecuada en una organización no es tarea sencilla. Es necesario conocer las necesidades que tienen los clientes para poder brindar un producto o servicio que cumpla sus necesidades. Es decir, entender la respuesta que la organización tiene que entregar basado en este comportamiento de los clientes.

Claramente, entre distintos sectores la complejidad del asunto puede cambiar. Uno de los sectores más complejos es el sector de la salud y dentro de éste, los actores críticos son los prestadores de salud, específicamente, los hospitales. Fuera de ser un servicio indispensable para subsistencia de la civilización, este tiene muchos factores que afectan la prestación que varía entre países, volviendo el estudio de este aún más complejo. Los aspectos distintivos de la prestación del servicio se ven desde los riesgos de salud de cada país, la morbilidad y mortalidad que padece la población, la disposición de recursos tecnológicos y financieros, y personal para la prestación del mismo. Junto con esto, se encuentra sumido en un mundo de relaciones, donde el servicio es afectado por factores políticos, económicos, sociales, demográficos y culturales (*SERVICIOS DE LA SALUD*, n.d.). Si a esto se le suma la alta demanda, la variabilidad y la incertidumbre asociada al perfil del cliente (Ageron, Benzidia, & Bourlakis, 2018), se empieza a evidenciar que la gestión logística hospitalaria requiere de gran análisis para poder llevarse a cabo de manera exitosa.

Según Choi, Dooley, & Rungtusanatham(2001) se puede encontrar una posible solución al problema observado de la dinámica compleja de la prestación del servicio hospitalario. Además, permitiendo estudiar los aspectos críticos mencionados previamente como perfiles socioeconómicos y demográficos (Rentería-Ramos & Vitale Alfonso, 2017). En cualquier lugar del mundo, en la prestación de este servicio, la carencia de calidad puede traer consecuencias tales como; esperas por la atención, cancelación de citas, el desabastecimiento de medicamentos, limitaciones en las instalaciones y recursos, y finalmente errores en el diagnóstico, lo cual puede llevar finalmente a casos de mortalidad específicos.

Colombia cuenta con un alto interés por parte del gobierno de incluir a todos sus ciudadanos en un programa de salud de alta calidad. Esto se puede observar con las leyes que ha establecido y los programas de salud con los que cuenta, dentro

de los cuales se encuentra la Ley Estatutaria de Salud, la cual busca la salud como derecho fundamental autónomo, garantizando su prestación y regulando esta. Algunos de los mayores retos que enfrenta esta ley es minimizar la desigualdad en el servicio, garantizar mayor aseguramiento y acceso a la salud y finalmente conseguir más resultados en términos de salud (MINSALUD, 2015). No solo se encuentra interés por parte de la legislación, también se evidencia que la cobertura del aseguramiento en salud representa un alto porcentaje de la población, específicamente para el 2018 esta cifra fue del 94.88% (MINSALUD, 2018). Pero la situación en lo que respecta la calidad de la atención no parece divergir de los problemas del sector, más aún dado el panorama complejo de desigualdad que tiene Colombia, la población vulnerable termina siendo la más afectada por estas situaciones de mala calidad. Como menciona Castro Hoyos (2012), en Colombia todavía se encuentran problemas para llegar a la universalidad y la eficiencia en el sistema de seguridad social en salud, este con un retraso cerca de una década. Además de esto, Castro Hoyos (2012), afirma lo siguiente, “ el fortalecimiento de la práctica de salud pública, la información completa y oportuna, así como la evaluación de sus procesos y resultados, aparece como un reto mayor sin el cual será difícil orientar el sistema...”. Es por esto que es de vital importancia establecer una gestión logística adecuada, para poder orientar el sistema para obtener buenos procesos y resultados.

Como mencionan Rentería-Ramos & Vitale Alfonso (2017), estos problemas de calidad no solo se ven en entidades públicas, también se observan en hospitales nivel III y nivel IV los cuales tienen los mejores servicios y los procesos con mayor complejidad. Además menciona que este problema se acentúa aún más en los hospitales públicos en donde la prestación del servicio es aún más compleja por el perfil de la demanda. Es evidente que es necesario buscar modelos de gestión logística hospitalaria que estén en capacidad de identificar las necesidades de los clientes y mejorar la calidad del servicio.

La tendencia a buscar un servicio médico de alta calidad no es solo mundial ni nacional, en la ciudad de Medellín en el Plan Municipal de Salud (Correa et al., 2012) se observan unos diagnósticos referentes al funcionamiento del sistema de salud. Donde se destacan aspectos a mejorar tales como: Mejorar la rectoría, favorecer la articulación entre actores, instituciones y programas y finalmente estabilizar la operación del régimen subsidiado. Uno de los principales actores en la salud de la ciudad de Medellín es el Hospital General de Medellín, un hospital el cual se destaca por ser hospital público de tercer nivel acreditado en salud nacional e internacionalmente, además con compromisos fuertes con la docencia y con la búsqueda hacia la seguridad de la prestación de servicio, motivo por el cual han establecido sistema de historial clínico electrónico.

Al tener un sistema rezagado de prestación de servicios hospitalarios en Colombia, un interés por parte de la ciudad de Medellín en mejorar la calidad de sus servicios médicos, una demanda sumida en un mundo complejo, la criticidad de la prestación del servicio, la falta de herramientas para estudiar la demanda por parte de las entidades y prestadores de servicios de salud pública es ¿Cómo es posible obtener un mayor conocimiento descriptivo de la demanda, para la gestión logística hospitalaria en hospitales nivel III en la ciudad de Medellín?

1.1.2. Justificación

La salud es uno de los pilares más importantes para el desarrollo y la equidad de cualquier país. Por lo que buscar servicios de salud de alta calidad se vuelve un aspecto inevitable para las políticas de cualquier país, específicamente en sus prestadores fundamentales de estos servicios. Como menciona (Volland, Fügner, Schoenfelder, & Brunner, 2017) en las últimas décadas, la logística se ha determinado como factor esencial para el manejo de los costos hospitalarios, sin embargo, a esta no se le ha dado la importancia que requiere en el sector. Esto se debe según a los autores a una creencia que considera que la reducción de costos va a llevar a la reducción de calidad de servicios. Sucedería lo contrario, con una adecuada gestión se reducirían costos y el tiempo invertido por los médicos y personal en el manejo de tareas logísticas, además se vería mejor invertido su tiempo en la atención de los pacientes con los insumos y flujos de información adecuados permitiendo así mejorar la calidad del servicio entregado.

Más aún, la importancia de definir un buen modelo de gestión logística hospitalaria se observa en las posibles oportunidades de mejora que se pueden obtener con procesos eficientes, en base a una buena identificación en los patrones de morbilidad de los pacientes. Según (Rolbina et al., 2018) el incremento en informatización y por ende, en análisis de datos tiene una correlación positiva con el incremento en la cobertura y calidad del servicio en salud. Por lo tanto es de vital importancia que se apliquen este tipo de estrategias de análisis para Colombia, en pro de mejorar la cobertura y calidad de los servicios de salud en el país.

La búsqueda por la calidad de la salud se puede evidenciar en los objetivos de desarrollo sostenible relacionados con la salud de las (Naciones Unidas, 2018). Uno de sus objetivos es el objetivo 3, el cual plantea lo siguiente, "Garantizar una vida sana y promover el bienestar para todos en todas las edades".

De donde se desprenden unas metas fundamentales tales como:

- 3.1. Para 2030, reducir la tasa mundial de mortalidad materna a menos de 70 por cada 100.000 nacidos vivos.
- 3.2. Para 2030, poner fin a las muertes evitables de recién nacidos y de niños menores de 5 años, logrando que todos los países intenten reducir la mortalidad

neonatal al menos hasta 12 por cada 1.000 nacidos vivos, y la mortalidad de niños menores de 5 años al menos hasta 25 por cada 1.000 nacidos vivos.

Otros de sus objetivos más relevantes para la región Latinoamericana y en especial Colombia se evidencian a continuación:

3. c. Aumentar sustancialmente la financiación de la salud y la contratación, el desarrollo, la capacitación y la retención del personal sanitario en los países en desarrollo, especialmente en los países menos adelantados y los pequeños Estados insulares en desarrollo.

3. d. Reforzar la capacidad de todos los países, en particular los países en desarrollo, en materia de alerta temprana, reducción de riesgos y gestión de los riesgos para la salud nacional y mundial.

Finalmente, más importantemente en su objetivo las (Naciones Unidas, 2018) busca la siguiente meta:

3.8 Lograr la cobertura sanitaria universal, en particular la protección contra los riesgos financieros, el acceso a servicios de salud esencial, de calidad y el acceso a medicamentos y vacuna seguros, eficaces, asequibles y de calidad para todos.

Este último objetivo es un punto de partida fundamental para buscar una gestión logística hospitalaria adecuada. Considerando los costos que se tienen por parte de los insumos, el efecto que tienen los insumos en la calidad de la prestación de servicio y la oportunidad de mejora se hace evidente que es necesario buscar un sistema de gestión logística hospitalaria basada en análisis profundo de datos que apunte a la dirección de estos objetivos. Solo en Colombia para el 2014 la tasa de mortalidad infantil reportada fue del 5.88%, bastante superior a la tasa sugerida por la OMS, en Medellín esta tasa fue del 8,1% en el 2015, mostrando también un rezago frente a los objetivos planteados.

Si los modelos convencionales no han logrado dar un buen aporte y según (Li, Tao, Cheng, Zhang, & Nee, 2017) los modelos logísticos son una sistema dinámico complejo y la optimización y control de estos modelos son un tema de investigación corriente. Se puede evidenciar con lo dicho en la formulación del problema que la red logística hospitalaria está sujeta a ser estudiada mucho más a profundidad. Más aún, está en orden identificar una mejor caracterización de la demanda en el modelo de gestión logística para hospitales con funciones relevantes y con alta demanda quirúrgica, como lo pueden ser los hospitales nivel III. En específico para políticas preventivas y preparación en un contexto, el que se vive en Medellín, que es una ciudad ubicada en un país en desarrollo, con una oportunidad drástica de mejora y un entorno socioeconómico complejo. Es acá donde importantes herramientas podrán jugar un papel clave en la mejora de los modelos de salud. Herramientas avanzadas como el RPA (automatización robótica de procesos) permiten agilizar el proceso de caracterización y llegar a un nivel alto de especificidad de vital

importancia para conocer el comportamiento de la morbilidad en el entorno. (Aguirre & Rodríguez, 2017) señalan esta herramienta como fundamental para automatizar procesos que serían repetitivos, estructurar datos y resultados determinísticos. Como resultado, se puede observar un incremento de la productividad por el uso del RPA. Una herramienta que además de ser eficaz permitiría ahorros de tiempo, que se podrían destinar a otras tareas de análisis profundos.

1.2 OBJETIVOS DEL PROYECTO

1.2.1 Objetivo General

Desarrollar una metodología para el análisis descriptivo de la demanda, basado en técnicas de análisis de datos, para hospitales nivel III en la ciudad de Medellín basada en los RIPS.

1.2.2 Objetivos Específicos

- Recolectar RIPS del HGM para el período 2014-2017
 - Realizar trámite de solicitud de datos
 - Solicitar datos de morbilidad y procedimientos para el Hospital General de Medellín (Hospital Nivel III).
- Estructurar códigos SQL y RPA para el análisis de datos
 - Limpiar y filtrar datos para ingresar al software de análisis
 - Realizar análisis descriptivo de los datos mediante el desarrollo de un robot en RPA
 - Realizar ABC de los procedimientos médicos para cada enfermedad crítica en la morbilidad
- Modelar en series de tiempo los procedimientos críticos
- Analizar resultados de la demanda
 - Buscar tendencias y estacionalidad
 - Buscar patrones en la morbilidad

1.3 MARCO DE REFERENCIA

1.1.3. Antecedentes

Los métodos analíticos de datos se han utilizado en varias aplicaciones con el fin de estudiar la cadena de suministro en aplicaciones hospitalarias. Varios autores recomiendan que se realicen estudios con herramientas cuantitativas más avanzadas, ya sea desde su mera definición hasta la aplicación de estas para estudiar las cadenas de suministro. Además, establecen la importancia de estudiar la morbilidad como mecanismo de prevención y encontrar nuevos patrones en los datos de demanda de estas cadenas logísticas.

- (Bernal, Forero et al., 2011) realizan una caracterización de la morbilidad en Colombia para el período 2004-2008. Estos mencionan la importancia que tienen los datos para el desarrollo de políticas públicas en el país. Recomendán además fortalecer este mecanismo de recolección y análisis ya que se menciona la pobre tarea y la poca disponibilidad de información que existe en el país. Estos concluyen que el país no tiene un “perfil de morbilidad claramente especificado”, por lo que se incita a investigación más a fondo.
- En su investigación sobre la logística de materiales en hospitales, (Volland, Fügner, Schoenfelder, & Brunner, 2017) identifican la gestión de materiales dentro de los hospitales como una palanca clave para contener los costos de la salud en los países industrializados. Además invitan a que se realice investigación más profunda respecto a aplicación de métodos cuantitativos en cuanto a la gestión holística de la cadena de suministro y que se realice investigación respecto a incluir el cuidado de la salud en la investigación de operaciones. Dando un pie de partida para investigar posibles mejoras aplicando modelos de análisis para la demanda en la gestión logística.
- (Volland et al., 2017) desarrollaron un estudio sobre los retos a los que se enfrenta la práctica hospitalaria sostenible en Dinamarca. En su estudio identificaron que la logística integrada era el paso a continuar para poder dar abasto con las necesidades cambiantes y crecientemente demandantes de la salud en Dinamarca. Según ellos los hospitales daneses tienen un gran foco en el desarrollo organizacional y también tecnológico, esto ofreciendo un gran potencial de innovación en el campo de logística hospitalaria. Esta investigación abre campo a más investigaciones respecto a la logística hospitalaria, en donde la innovación permite buscar aplicar teorías de gran relevancia como lo pueden ser los métodos de análisis de datos.
- En un estudio realizado para la República de Tatarstan, (Rolbina et al., 2018) identifican como la informatización ha tenido un impacto directo positivo en la calidad de los servicios de salud en el país. Presentan y proponen mejoras

al sistema de información de EEUU para mejorar la calidad del servicio que se ofrece en el país. Este estudio podría ser aplicado para Colombia en pro de mejorar los sistemas de información de salud para posteriormente mejorar la prestación de los servicios.

- En el caso de (Willcocks, 2017) se presenta la herramienta de RPA para aplicaciones en servicios con operaciones globales. La experiencia con una de las aplicaciones de la herramienta, Xchanging, en el mercado de seguros en Londres fue bastante satisfactoria. Los autores realizan una capacitación por estudiantes y sugieren la aplicación de la herramienta en diferentes áreas de prestación de servicios. Además, se plantean la pregunta de si esta herramienta puede ser utilizada únicamente para decisiones explícitas (tareas automatizables) o si en el futuro podrá ser también utilizada con fines estratégicos.

1.1.4. Marco Teórico

Se define primero el concepto holístico del análisis de datos como el proceso sistemático de aplicar técnicas estadísticas o lógicas para describir y evaluar datos. En el proceso se inspecciona, limpia, transforma y modela una serie o series de datos con el objetivo de descubrir información relevante, extraer conclusiones o ayudar en la toma de decisiones. Según los autores (Xia, Gong et al., 2015), el análisis de datos juega un papel bastante importante en la toma de decisiones de forma “más científica” y ayuda a las empresas y entidades públicas a desarrollar procesos cada vez más eficientes.

El análisis de datos como concepto macro, se presenta con cantidades de técnicas y facetas usando muchísimos nombres, todas con el mismo fin de descubrir y describir fenómenos para mejorar la toma de decisiones.

SISPRO: El proceso SISPRO permite, por medio del Sistema de Gestión de datos (SGD) “obtener, procesar, consolidar y proveer a los diferentes tipos de usuarios, de la información necesaria para la toma de decisiones que apoye la elaboración de políticas, el monitoreo regulatorio y la gestión de servicios, en cada uno de los niveles del sector de la protección social suministrando información a todos los usuarios y a la población en general... La Bodega de Datos de SISPRO le permite al usuario consultar y elaborar reportes a partir de una construcción dinámica que se obtiene con la navegación de los elementos fundamentales de una bodega de datos: dimensiones y medidas.

El SISPRO es una bodega de datos compuesta por datos tanto internos como externos del Ministerio de Salud. Este se compone de los siguientes componentes: RUAF, RIPS, SIHO, SGD, SISMED entre otros.” (MinSalud, 2013)

El acceso a los datos se logra mediante el Sistema de Gestión de datos, el cual como está compuesto por unas bodegas de datos, como lo son el cubo de los RIPS en este caso.

RIPS

Los datos son Los Registros Individuales de Prestación de Servicios de Salud – RIPS. Los datos son definidos como “el conjunto de datos mínimos y básicos que el Sistema General de Seguridad Social en salud requiere para los procesos de dirección, regulación y control y como soporte de la venta de servicios, cuya denominación, estructura y características se ha unificado y estandarizado para todas las entidades a que hace referencia el artículo segundo de la resolución 3374 de 2000 (IPS, profesionales independientes, grupos de práctica profesional, entidades administradoras de planes de beneficios y los organismos de dirección, vigilancia y control SGSSS.)” (MinSalud, 2015)

Los datos están caracterizados por cuatro clases: Datos de identificación, datos del servicio, datos del motivo de la atención y datos del valor. Estos datos están sujetos a un validador que proporciona el Ministerio, el cual hace una verificación para comprobar que no tengan errores.

Los datos poseen tanto variables cualitativas (tipo de enfermedad, diagnóstico, causa, etc...) como cuantitativas (costo del procedimiento, fecha, edad, etc...). Se categorizan en dos secciones: Variables de consulta del CUBO de los RIPS y hechos o medidas del CUBO de los RIPS.

Fecha de atención	Se refiere a la variable tiempo que corresponde a la fecha de la atención en la prestación del servicio de salud por consulta, hospitalización, procedimientos o urgencias.
Ambito de Procedimiento	Identifica el ámbito de realización del procedimiento según la ubicación funcional donde éste se realiza. Ver Anexo Dominios para dimensiones.
Diagnóstico Principal y Diagnóstico Principal Egreso	Motivo que origina la prestación del servicio. Código del diagnóstico principal, según la Clasificación Internacional de Enfermedades vigente. Se debe tener en cuenta que la fuente diligencia este campo para procedimientos quirúrgicos y para los egresos.
Finalidad de procedimiento	Identifica la finalidad con que se realiza el procedimiento. Ver Anexo Dominios para dimensiones.
Grupo Etéreo	Permite consultar las diferentes medidas utilizando los rangos de edades para diferentes tipos de grupos etéreos, ver Anexo

Tabla 1: Ejemplo de algunas variables cualitativas y campos de los RIPS. Obtenido de: “Documentación Reportes RIPS V4” (Min Salud, 2013)

Grupo de medidas	Medidas	Descripción
Conteo de Prestadores	Conteo de Prestadores	Corresponde al número de prestadores por los que se reportó información en el período consultado.
Personas Atendidas	Número de Personas Atendidas	Corresponde al número de personas únicas atendidas. Se cuenta una sola vez el usuario, independiente si ha recibido más de una atención. Las personas únicas se identifican con tipo y número de documento de identificación.
Prestación de Servicios de Salud	Número de Atenciones	Corresponde al número de atenciones prestadas y reportadas en el periodo consultado.
	Número de Días de Estancia	Hace referencia a la sumatoria del número de días de estancia reportados en el archivo de Hospitalizaciones.
Valores	Concentración	$\text{Número de atenciones} / \text{Número de personas atendidas}$
	Porcentaje de usuarios	$\text{Número de personas atendidas} / \text{Número total de personas atendidas}$
	Promedio días de estancia	$\text{Sumatoria día de estancia} / \text{Número de personas atendidas}$

Tabla 2: Variables cuantitativas de los RIPS. Obtenido de: “Documentación Reportes RIPS V4” (Min Salud, 2013)

A continuación, se desarrolla el pre tratamiento de datos, lo cual es definido en términos técnicos como “data cleaning and cleansing.” Este proceso se encarga de detectar, corregir o borrar todos los datos corruptos o incorrectos de los datos que van a ser usados en los análisis (Wu S., 2013). Esta etapa garantiza que los datos que se usarán podrán ser utilizados con precisión y tengan un formato compatible con las herramientas de análisis que se utilizarán después. El pre tratamiento de datos también puede ser realizado mediante muchas técnicas y herramientas, dos de las más conocidas son el “data wrangling;” donde se transforman datos crudos de un formato al otro y procesamiento en batch, donde se aplica una secuencia de pasos para procesar toda la serie de datos. Uno de los lenguajes más conocidos para realizar este procesamiento en batch es SQL.

SQL: El término SQL hace referencia a “Structured Query Language” por sus siglas en inglés. Es un lenguaje de programación que permite comunicarse con las bases de datos y es comúnmente utilizado para relacionar bases de datos. Permite acceder, modificar y crear datos en bases de datos. Se utilizará para el estudio la interfaz del lenguaje de programación mediante la herramienta “Power Query” en Excel. Esta herramienta permitirá realizar una serie de filtros, transformaciones, eliminaciones y agrupaciones en masa, por los cuales pasarán los datos extraídos RIPS, para posteriormente ser utilizados en herramientas de análisis como RPA.

RPA o Robotic Process Automation por sus siglas en inglés. Esta herramienta permite la programación de un software o “robot”, para realizar operaciones que normalmente serían llevadas a cabo por un humano, comúnmente en aplicaciones de negocios. RPA captura datos y manipula aplicaciones como cualquier ser humano y lleva a cabo una gran variedad de tareas que serían repetitivas.

El RPA se utiliza con el fin de automatizar tareas repetitivas y tediosas basadas en reglas de lógica de procesos. Estos funcionan con base en unas entradas básicas y son efectivos a la hora de procesar grandes cantidades de datos, funcionar entre varios aplicativos, reducir el error humano y reducir costes en las empresas.

Además de presentar un beneficio administrativo, la automatización robotizada de procesos (RPA) permite que los encargados de procesos puedan dedicarles más tiempo a las tareas de análisis, donde el tiempo empleado genera mayor valor.

El RPA se ha utilizado en una gran cantidad de industrias como lo son las industrias; financiera, logística, salud, servicios y recursos humanos.

UiPath es una plataforma abierta de desarrollo y gestión de automatización robótica de procesos RPA. Esta cuenta con varios productos los cuales son:

Studio: Este producto es donde se lleva a cabo el desarrollo del RPA. En este se tiene la posibilidad de utilizar varios lenguajes de programación con múltiples actividades predeterminadas las cuales pueden ser arrastradas mediante un “drag and drop” a un diagrama de flujo o secuencia de actividades.

Robot: Robot es donde se realiza la compilación del RPA desarrollado en “studio”.

Orchestrator: Finalmente se encuentra “orchestrator”, que como su nombre indica es una herramienta para determinar correr robots programados, establecer horarios, colas e indicadores.

En este caso se eligió el software UiPath ya que tiene una versión de comunidad que es gratuita, la cual tiene las mismas funcionalidades que la versión paga siempre y cuando no se utilice para fines comerciales o lucrativos.

Con el fin de entender el desarrollo del estudio de la demanda es preciso conocer algunos términos claves de la plataforma de UiPath:

Existen dos tipos de formatos de orden lógico de actividades. El primero es una secuencia, en donde las actividades van dentro de una casilla en formato lineal una tras otra. Por otro lado, está el diagrama de flujo o “Flowchart” el cual tiene el mismo aspecto físico que un diagrama de flujo convencional solo que se compone por actividades predeterminadas. Las secuencias son más comúnmente usadas para flujos lineales no complejos con pocas decisiones, por el contrario, los diagramas

de flujo se utilizan más comúnmente para reglas de negocio complejas o múltiples decisiones y actividades largas dentro de un proceso. En este caso también entra a jugar un papel la estética y los estilos de desarrollo personal.

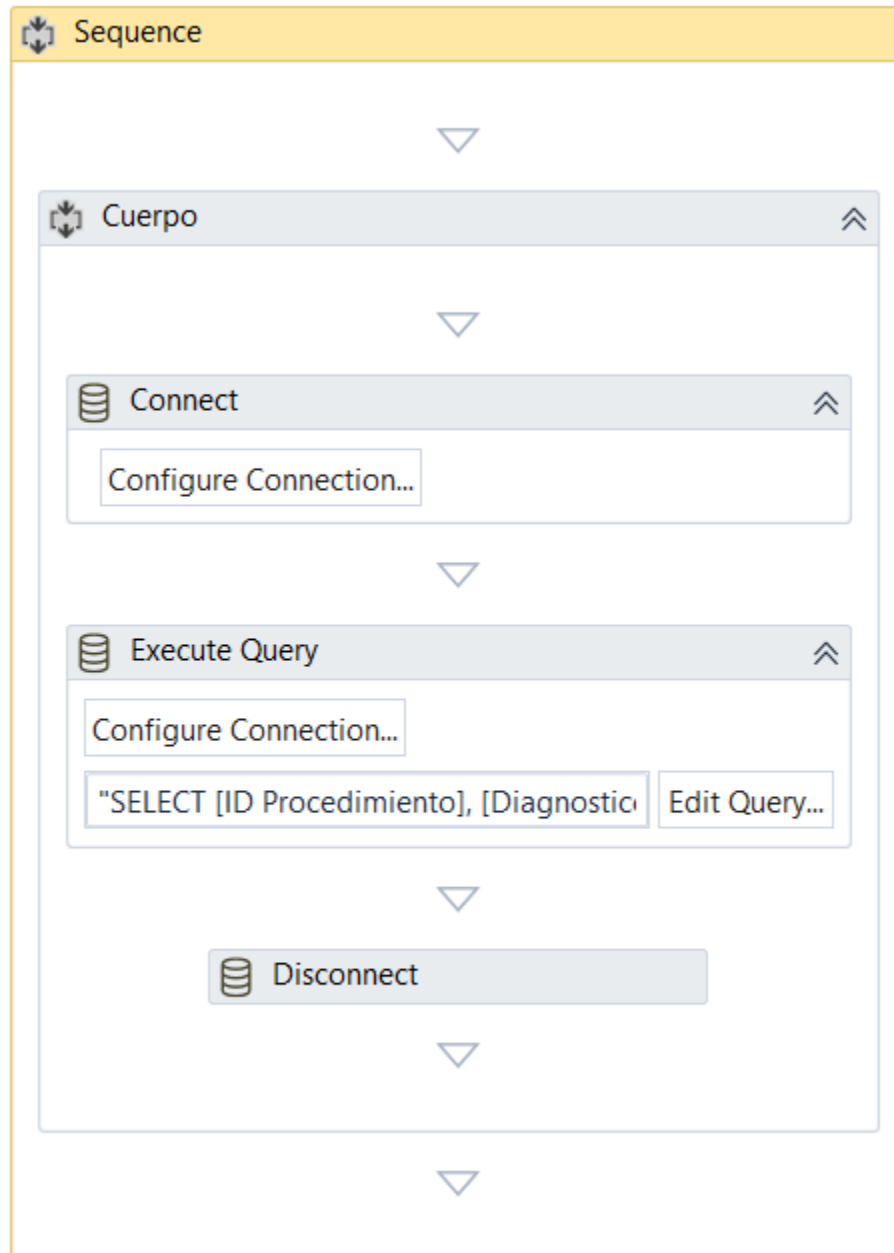


Figura 1: Imagen de una secuencia en UiPath. Obtenido de: Autoría propia

Una secuencia lineal mostrando además unas funciones de SQL.

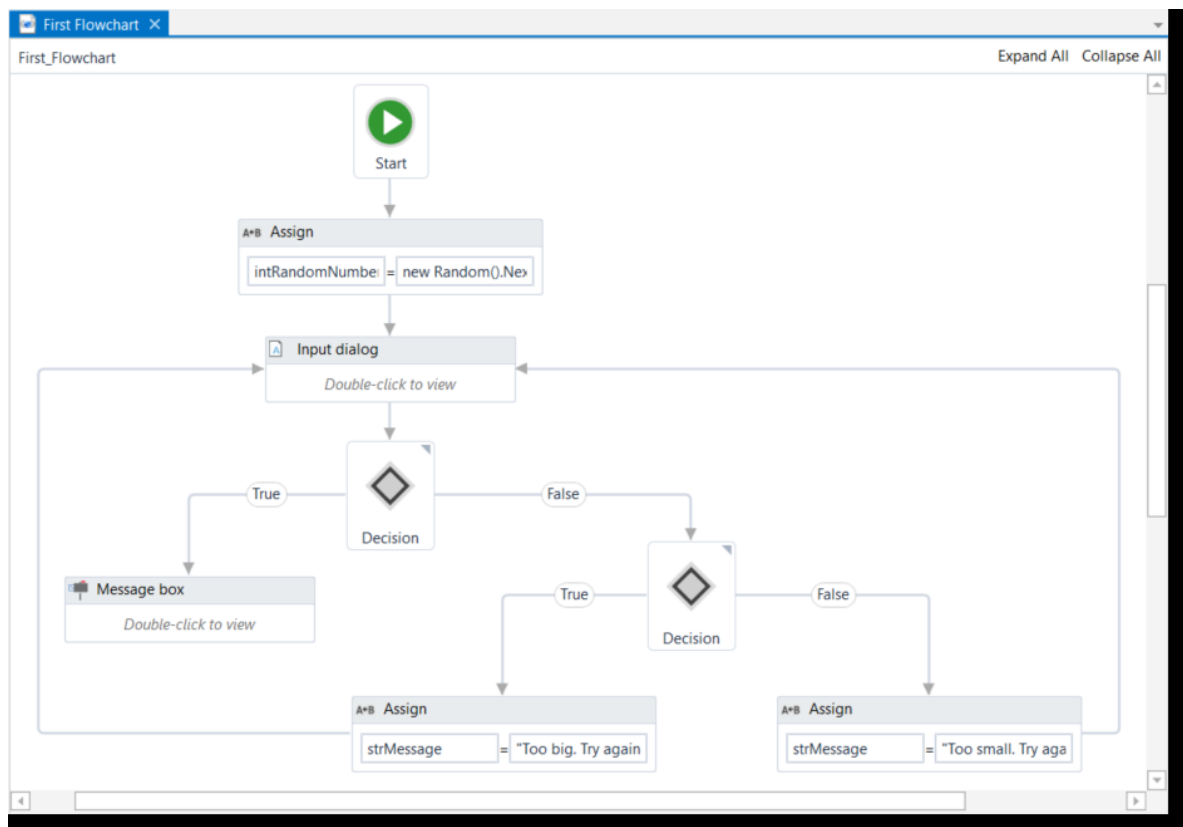


Figura 2: Imagen que muestra el esquema de un flowchart en UiPath. Fuente: Autoría propia

Actividades:

Dentro de las actividades más comunes y de relevancia para esta investigación se encuentran las siguientes:

1. Assign: Esta actividad se utiliza para asignar valores de variables.
2. Read range: Esta actividad permite leer un archivo de Excel únicamente con valores generando como resultado una tabla de datos para manipular, analizar y transformar en UiPath.
3. Write Range: Similar a la actividad anterior permite escribir una hoja de Excel con base en una tabla de datos predeterminada.
4. For each row: Esta actividad es única y especialmente para tablas de datos, es un método iterativo que recorre la tabla de datos fila por fila. Cada fila que recorre la transforma en una variable tipo row temporalmente para permitir manipulación de los datos.
5. Get row item: El get row item se utiliza para extraer información de una variable de tipo fila, identificando el valor a sacar por medio de unos parámetros configurables como nombre de columna, índice de columna etc. El get row item se encuentra comúnmente dentro de un for each row.

6. Add data row: Permite ingresar un arreglo de datos en formato de fila a una tabla de datos, ya sea mediante un arreglo o una secuencia de variables.
7. Add data column: Similar al Add data row permite añadir una columna a una tabla de datos.
8. Build data table: Se utiliza para crear una tabla de datos dentro de la plataforma, asignando columnas, variables para cada columna y nombres de cada columna.
9. Sort data table: Es usado para organizar una tabla de datos mediante el nombre o identificador de columna y tiene opciones de ordenado ascendente o descendente.
10. Filter data table: Filtra una tabla de datos con base a unas condiciones ingresadas.
11. IF: La actividad del if es un condicional lógico “si”. Es utilizado mayormente para secuencias ya que para flujogramas se tiene la actividad decisión de flujo.
12. Connect: Forma parte de un paquete de actividades de bases de datos. Este genera una conexión a una base de datos para luego incorporarle otras actividades.
13. Query: Se basa en la conexión creada por el connect para hacer consultas sobre tablas de datos mediante el lenguaje SQL en Ui Path.
14. Disconnect: Remueve la conexión creada de la actividad connect.

Clasificación ABC:

La clasificación ABC es tradicionalmente utilizada en ingeniería industrial para controlar y clasificar productos (SKU) que pertenecen a una misma categoría. “La clasificación de los productos en categorías se realiza de acuerdo a su importancia o relevancia por algún tipo de criterio...” (Castro-Zuluaga & Gallego, 2011). Normalmente, la clasificación se hace en base a un criterio de Pareto, en el cual un 20% del número de los SKU's representa un 80% del criterio (costo, por ejemplo). Este 20% obtiene una clasificación “A”, los cuales serán los productos más relevantes y los cuales tradicionalmente se analizan con mayor rigurosidad. En el caso particular, serán materia de estudio solo los procedimientos que obtengan una clasificación en esta categoría.

Series de Tiempo:

Una serie de tiempo es un conjunto o sucesión de variables aleatorias ordenadas por una unidad de tiempo. Las series de tiempo permiten visualizar el comportamiento de los datos presentes para tratar de predecir o identificar futuros

comportamientos, con el supuesto de que los datos se comportarán de la misma forma en el futuro.

Las series de tiempo se pueden analizar descomponiéndolas en tres partes:

- Tendencia: Muestra el crecimiento o decrecimiento de los valores de la serie. Es la evolución de los datos en el largo plazo.
- Estacionalidad: Muestra un patrón que tiene un cambio regular durante un mismo período en el tiempo, normalmente se repite con una base anual. Es decir, los datos tienen el mismo comportamiento durante fechas específicas en diferentes períodos de tiempo. Es un comportamiento periódico de los datos.
- Error: Variaciones aleatorias que no siguen ningún comportamiento definido. Muchas veces es referido como “ruido” dentro de los datos.

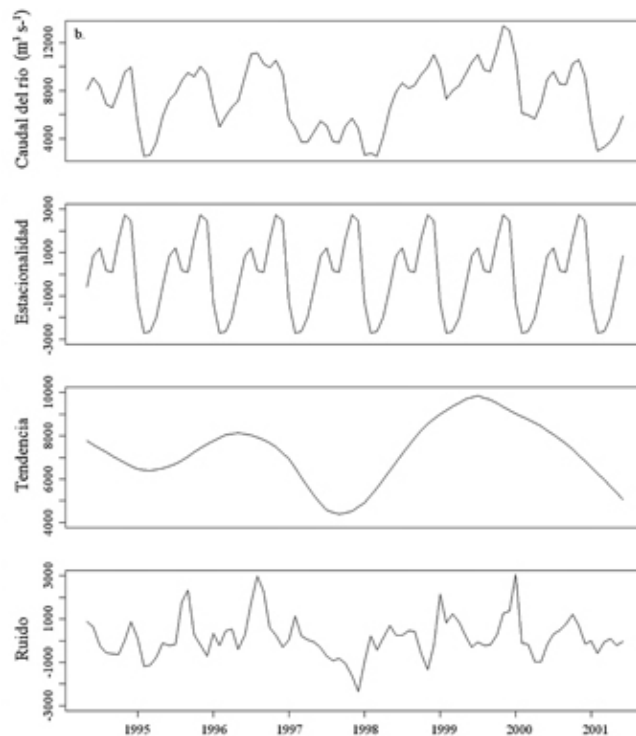


Figura 3: La primera gráfica, titulada “Caudal del río” muestra la serie de datos graficados en el tiempo. La segunda, muestra la componente estacional de esos datos. La tercera es la componente de tendencia de los datos y la última, es el error. Obtenido de: (Párraga, 2010) en: https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0718-19572010000100007

Existen, además, dos métodos de descomposición en las series de tiempo. El primero es el modelo aditivo. Este se adecua cuando la estacionalidad de los datos no crece con la tendencia. El segundo modelo, el modelo multiplicativo, se

aplica cuando el efecto estacional incrementa o decrece obedeciendo a la tendencia.

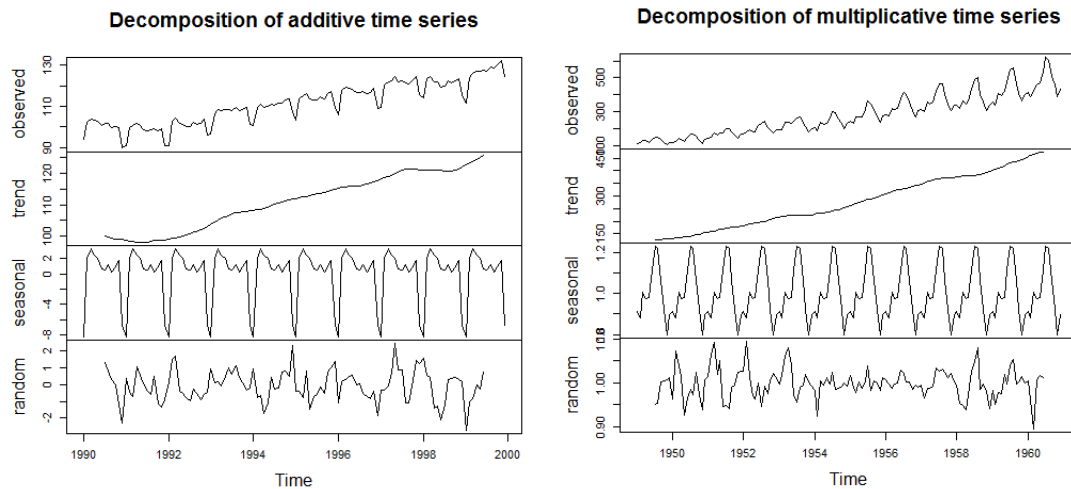


Figura 4: A la izquierda de la imagen se puede observar una descomposición de una serie de tiempo con descomposición aditiva; se observa como la estacionalidad no crece con la tendencia. Del otro lado, vemos una descomposición de una serie multiplicativa, donde se observa el crecimiento del patrón estacional con la tendencia. Obtenido de: Notas de clase Laura Ardila, Procesos Estocásticos, 2018.

Métodos para aislar las componentes de Tendencia y Estacionalidad:

Existen dos tipos de métodos para aislar las componentes de una serie de tiempo. Los métodos NO paramétricos son aquellos que aíslan las componentes utilizando la “suavización” de la serie de datos. Dentro de los métodos no paramétricos se encuentran los modelos de medias móviles, medias estacionales, tendencias lineales, aislado exponencial de Holt, suavizamiento Holt-Winters, entre otros. Por otra parte, los métodos paramétricos aíslan las componentes mediante modelos ajustados, estos “proponen modelos paramétricos para expresar la relación que guardan la tendencia y la componente estacional con el tiempo” (Aneiros, 2008). Dentro de estos métodos encontramos las funciones de autocovarianza y autocorrelación, los modelos autoregresivos AR, los modelos de medias móviles MA y los modelos ARMA y ARIMA.

Función “decompose”:

La función decompose o descomponer, en R, funciona con la técnica de medias móviles para aislar las componentes de tendencia y estacionalidad, dando a la función como entrada si la serie de tiempo es multiplicativa o aditiva.

Finalmente, con el fin de interpretar las implicaciones financieras de la demanda se decidió realizar el Cálculo del capital regulatorio (OpVar) mediante simulación Monte Carlo:

Con el fin de calcular el capital requerido en respuesta al Riesgo Operativo (RO) el Comité de Supervisión Bancaria de Basilea II planteó una metodología con base en el enfoque de la distribución por pérdidas agregadas (LAD). Este enfoque de distribución de pérdidas agregadas se basa en el supuesto que las pérdidas son compuestas la frecuencia de las observaciones la severidad o el costo implicado. Conociendo estas dos entradas, para el cálculo del capital regulatorio se utiliza el concepto Valor en Riesgo o VaR aplicado al riesgo operacional tomando entonces el nombre OpVar (Operational Value at Risk)(Flores, 2012). Este concepto es una representación de un percentil de las pérdidas generadas debidas al RO. Generalmente este consiste en estimar la función de pérdidas y calcular al percentil superior de esta función, el cual indica el capital requerido para tener solvencia.

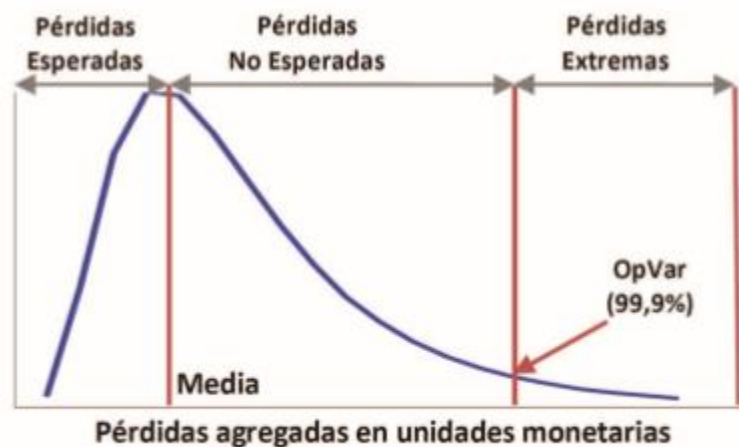


Figura 5: Imagen de la interpretación del OpVar y distribución de Pérdidas agregadas. Obtenido de: (Flores, 2012)

Con el fin de determinar el OpVar, se debe definir la distribución de pérdidas con las distribuciones de la frecuencia y la severidad.

La distribución de pérdidas se define como el producto entre la frecuencia y la severidad, siendo la frecuencia una referencia de la cantidad de eventos en un intervalo de tiempo, mientras la severidad representa la magnitud de la pérdida en el mismo horizonte de tiempo.

$$St = \sum_{k=1} NtXk \quad \text{Ecuación 1. LPA}$$

Siendo X_k la magnitud de la pérdida y N_t la distribución del conteo.

Generalmente X_k sigue una distribución de probabilidad continua y N_t sigue una distribución discreta. Dentro de las distribuciones recomendadas para la frecuencia se encuentran las distribuciones Poisson, Geométrica, Binomial Negativa y la binomial. Para modelar la severidad se recomiendan distribuciones como la Lognormal, la Pareto, la Rayleigh, Weibull y la Exponencial.



Figura 6: Representación gráfica del LAD. Obtenido de: (Flores, 2012)

Teniendo definidas las distribuciones de probabilidad de los dos eventos causales y la definición distribución de pérdidas, para calcular el OpVar se propone utilizar simulación Monte Carlo en donde se estima mediante n simulaciones aleatorias la distribución de salida del LAD. Para un resultado adecuado es necesario definir la cantidad N de simulaciones que arroje la precisión deseada del modelo a analizar.

Finalmente, con la distribución (función de incertidumbre) de salida como resultado de la simulación se procede a calcular el valor del capital necesario a un nivel $(1-\alpha)$ para responder por estos casos de pérdidas extremas como manifestación del RO en los extremos de las colas de la distribución de pérdidas.

2. METODOLOGÍA

Pre-tratamiento de Datos:

En primera medida, se establece una conexión con la base de datos SISPRO, la cual contiene los RIPS. Esta conexión se realiza mediante Excel (connect to Server Analysis SQL). La conexión a la base de datos muestra los datos mediante tablas dinámicas.

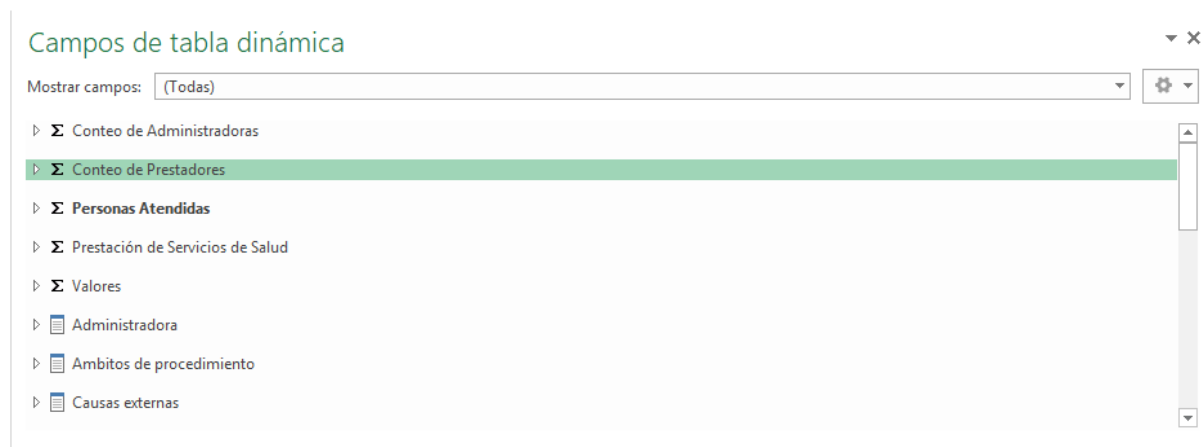


Figura 7: Ejemplo de una porción de los campos obtenidos de la base de datos SISPRO, mediante tablas dinámicas. Fuente: autoría propia basada en datos RIPS

Para extraer los datos se aplica un filtro por año y entidad prestadora del servicio (se debe garantizar que los datos provengan de hospitales nivel III), seleccionando año por año desde 2014 hasta 2017 y el Hospital General de Medellín como muestra para los hospitales nivel III, debido a que la cantidad de datos superan la capacidad de procesamiento de los computadores utilizados para el estudio. Se extraen los datos “raw” aplicando únicamente el filtro de año. Se obtienen 4 tablas de datos.

[Prestación de Servicios de Salud].[\$Procedimientos médicos.ID Procedimientos]	1
[Prestación de Servicios de Salud].[\$Fecha de Atención.Fecha]	2014/06/09
[Prestación de Servicios de Salud].[\$Personas.Persona Basica ID]	746073
[Prestación de Servicios de Salud].[\$Administradora.Administradora]	13-24 - LA PREVISORA S.A. COMPAÑIA DE SEGUROS

[Prestación de Servicios de Salud].[\$Diagnóstico Principal.Diagnósticos]	S898 - OTROS TRAUMATISMOS DE LA PIERNA, ESPECIFICADOS
[Prestación de Servicios de Salud].[\$Diagnóstico Principal Egreso.Diagnósticos]	0 - NO REPORTADO
[Prestación de Servicios de Salud].[\$Ambitos de procedimiento.Ambito del Procedimiento]	N - NO APLICA
[Prestación de Servicios de Salud].[\$Finalidad de procedimientos.Finalidad Procedimientos]	6 - NO APLICA
[Prestación de Servicios de Salud].[\$Finalidad Consulta.Finalidad Consulta]	10 - NO APLICA
[Prestación de Servicios de Salud].[\$Tipo Usuario.Tipo de Usuario]	1 – CONTRIBUTIVO
[Prestación de Servicios de Salud].[\$Personal que atiende.Personal que atiende]	0 - NO APLICA
[Prestación de Servicios de Salud].[\$Causas externas.Causa Externa]	02 - ACCIDENTE DE TRÁNSITO
[Prestación de Servicios de Salud].[\$Prestadores Basicos.Prestador Basico ID]	0500102144 - HOSPITAL GENERAL DE MEDELLIN LUZ CASTRO DE GUTIERREZ, EMPRESA SOCIAL DEL ESTADO
[Prestación de Servicios de Salud].[\$Tipo de Atencion.Tipo de Atención]	3 – URGENCIAS
[Prestación de Servicios de Salud].[\$Tipo Egreso.Tipo Egreso]	1 - NO DEFINIDO
[Prestación de Servicios de Salud].[\$Municipio Residencia - RIPS.ID Geografia]	12086
[Prestación de Servicios de Salud].[\$Tipo Diagnóstico Principal.Tipo Diagnóstico Principal]	ND - NO DEFINIDO
[Prestación de Servicios de Salud].[\$Estado Salida.Condicion Final]	1 - VIVO (A)
[Prestación de Servicios de Salud].[\$Edad de la Atención.Edad]	33
[Prestación de Servicios de Salud].[Número Días de Estancia]	0
[Prestación de Servicios de Salud].[Costo Procedimiento]	0
[Prestación de Servicios de Salud].[Valor Consulta]	0
[Prestación de Servicios de Salud].[Valor Cuota Moderadora]	0
[Prestación de Servicios de Salud].[Neto A Pagar - Consulta]	0
[Prestación de Servicios de Salud].[Valor Cuota Moderadora - Consulta]	0
[Prestación de Servicios de Salud].[Número de Atenciones]	1

Tabla 3: Tabla ejemplo (Nombre y primera columna) de los datos extraídos desde la base de datos, datos RIPS en sus formatos originales, sin modificaciones. Fuente: Datos RIPS, autoría propia

Mediante la herramienta “Power Query” de Excel se realiza la transformación de los datos mediante Queries, los cuales utilizan el lenguaje de programación de SQL para transformar de manera sistemática los datos.

Se realizan las siguientes transformaciones a los datos originales:

- Se renombran las columnas para cambiar los formatos impuestos por el lenguaje de programación.
- Se cambia el formato a la columna “Fecha”, pasando de un formato dd/mm/aaaa a indicar solamente el número de la semana del año (entre 1 y 52).
- Se aplica un filtro a la columna “Diagnóstico Principal” donde se eliminan todos los registros los cuales tengan un diagnóstico ingresado como “1 – No Definido.” De esta forma se eliminan registros de diagnósticos no relevantes o redundantes para la investigación.
- Se aplica un segundo filtro a la misma columna de “Diagnóstico Principal”, donde se eliminan los registros con el campo “Z768 – Persona en contacto con los servicios de salud en otras circunstancias especificadas” ya que se establece este dato como no relevante para la investigación pues no brinda información sobre el diagnóstico de una enfermedad.
- Se eliminan las siguientes columnas ya que no son necesarias en la investigación, con el fin de aligerar el volumen de datos para análisis:
 - “Tipo de Egreso”
 - “Prestadores Básicos”
 - “Finalidad Consulta” (Todos los campos ingresados con “No Aplica”)
 - “Tipo Diagnóstico Principal”
 - “Valor Cuota Moderadora” (Todos los campos ingresados con 0)
 - “Neto a pagar – Consulta”

Se aplica el mismo código SQL a las 4 tablas de datos diferentes, dando a estas el mismo pre tratamiento de datos.

Después de aplicar exactamente el mismo tratamiento a los datos extraídos de la base de datos que contienen los datos RIPS, se utiliza otro pedazo de código en SQL para unir las 4 tablas en una sola, unificando los datos en un solo archivo. Se realiza toda la operación mediante el código en el Anexo 1.

Se obtiene el resultado del pretratamiento de datos como una tabla en Excel con 865.217 entradas de datos.

ID Procedimientos	1
Semana	24
ID Persona	746073
Administradora	13-24 - LA PREVISORA S.A. COMPAÑIA DE SEGUROS
Diagnostico	S898 - OTROS TRAUMATISMOS DE LA PIERNA, ESPECIFICADOS
Diagnostico Egreso	0 - NO REPORTADO
Ambito de Procedimiento	N - NO APLICA
Finalidad Procedimiento	6 - NO APLICA
Tipo Usuario	1 – CONTRIBUTIVO
Personal Que Atiende	0 - NO APLICA
Causas Externas	02 - ACCIDENTE DE TRÁNSITO
Tipo de Atencion	3 – URGENCIAS
Municipio Residencia	12086
Estado Salida	1 - VIVO (A)
Edad	33
Días Estancia	0
Costo Procedimiento	0
Valor Consulta	0
Numero Atenciones	1
Genero	F
Month	6

Year	2014
------	------

Tabla 4: Tabla ejemplo (Nombre y primera columna) de datos filtrados y modificados mediante SQL en la herramienta Power Query. Son los datos que luego se alimentan RPA. Fuente: autoría propia basada en datos RIPS

Tamaño de Muestra:

Se establece el tamaño poblacional como 52.993.748 como el número de personas atendidas en hospitales de nivel III para la ciudad de Medellín, para las cuales se tiene un registro de morbilidad. Este número es obtenido de los mismos registros RIPS mediante dos filtros:

- Se establece un filtro para la columna de “Geografía Prestador” únicamente como Medellín.
- Se establece un filtro para la columna “Nivel del Prestador” únicamente como Nivel III.
- El período muestra se mantiene constante (año 2014-2017)

A continuación, se determina la varianza poblacional para la morbilidad, utilizando la frecuencia por enfermedad en la muestra filtrada anteriormente. Los resultados son los siguientes:

Varianza Poblacional	1,284,148,077
Desviación Estándar Poblacional	35,835

Para determinar el margen de error, es decir, el nivel de precisión estadística o el rango (de error) en el cual la proporción de la población está estimada, utilizando 865.217 entradas de datos, se sigue el siguiente procedimiento:

- Se utiliza la fórmula: (WW, D., 1999)

$$n = \frac{(N * X)}{(X + N - 1)}$$

Para:

- n = tamaño de la muestra = 865.217
- N = tamaño de la población = 52.993.748

$$X = \frac{(Z\alpha/2)^2 * p * (1 - p)}{MOE^2}$$

- $Z\alpha/2$ para un nivel de confianza del 99.99% = 3.8906
- p es la proporción en la muestra
- Se dan 3 valores a p debido a las varias proporciones que existen en la muestra, garantizando que el tamaño de muestra es preciso para todas las proporciones del estudio, se despeja el margen de error para confirmar la precisión de la muestra respecto a la población total:

p	MOE
20%	0.1659%
1%	0.0413%
0.001%	0.0013%

- Se confirma la precisión de la muestra y propósito del estudio, garantizando precisiones bastante altas sin importar la proporción real de la muestra. La precisión de la muestra será siempre mayor al 99%.

Clasificación de la Morbilidad:

Luego de comprobar la relevancia estadística de los datos, se procede a una caracterización de la morbilidad. Es decir, se estudia mediante los datos RIPS, las enfermedades que padecen las personas que son estudiadas.

Mediante un código en RPA, se obtiene la clasificación de las enfermedades más recurrentes durante el período estudiado y se clasifica la morbilidad mediante los grupos vitales, estos son, según el Ministerio de Salud (Min Salud, 2019): “primera infancia (0-5 años), infancia (6 - 11 años), adolescencia (12-18 años), juventud (18 - 26 años), adultez (27 - 59 años) y vejez (60 años y más)”. Se establece la frecuencia de las enfermedades, realizando una clusterización mediante cada uno de los grupos y separando la muestra por sexo, el cual viene nombrado como Género en los RIPS. Las características de cada grupo se observan en la siguiente tabla:

	Género	Edad
Grupo 1	Masculino	Primera infancia
Grupo2	Femenino	Primera infancia
Grupo 3	Masculino	Infancia
Grupo 4	Femenino	Infancia
Grupo 5	Masculino	Adolescencia
Grupo 6	Femenino	Adolescencia
Grupo 7	Masculino	Juventud
Grupo 8	Femenino	Juventud
Grupo 9	Masculino	Adultez
Grupo 10	Femenino	Adultez
Grupo 11	Masculino	Vejez
Grupo 12	Femenino	Vejez

Tabla 5: Tabla que muestra la composición de los diferentes grupos para la clasificación de morbilidad basada en género y etapas del ciclo vital. Fuente: Autoría propia

Segmentación General de la Población:

Etapas 1: Caracterización de morbilidad general

La caracterización de morbilidad general se generó mediante una secuencia en UiPath, que tiene como finalidad generar tres archivos finales. Estos siendo sobre el resultado de la morbilidad, el resumen de morbilidad y finalmente un archivo de validación. El resultado de la morbilidad contiene la frecuencia y porcentaje asignada de diagnósticos para la muestra elegida, el resumen, como se indica contiene únicamente los primeros dos diagnósticos más frecuentes elegidos para posterior análisis y finalmente el resultado de validación contiene los conteos realizados para verificar que la lógica empleada estuviera correcta.

Descripción Caracterización de Morbilidad General con RPA:

Las actividades de la Caracterización de Morbilidad General son las siguientes:

1. Crear tabla de datos resumen con campos TotalRegistros y SumaEnfermedades mediante una actividad Build Data table.
2. Leer DatosRIPS mediante una actividad Read Range.
3. Ordenar la tabla de datos del literal 1 por el campo “Enfermedades” mediante un Sort Data Table.
4. Asignar la variable totalRegistros con el valor de la longitud de la tabla de datos mediante la actividad Assign.

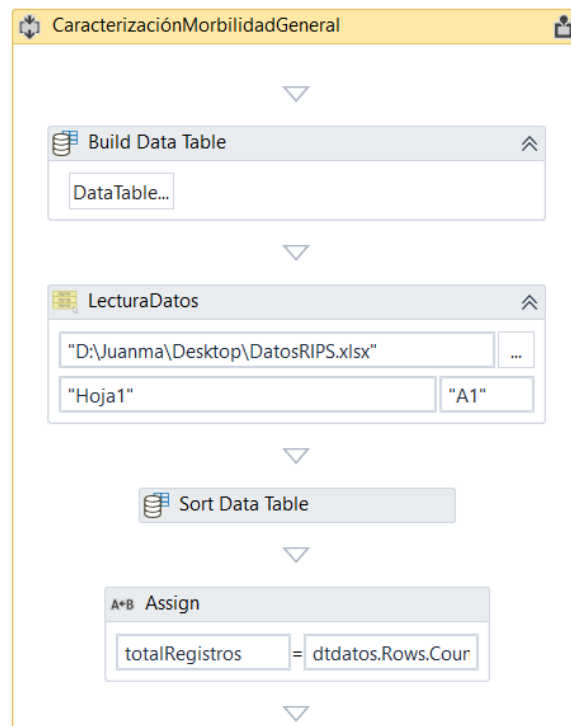


Figura 8: Actividades 1, 2, 3 y 4 de la Caracterización de Morbilidad General. Fuente: Autoría propia

5. Recorrer la tabla de datos ordenada en el literal 3 mediante una actividad For each row, establecer condiciones lógicas y acumular variables como sigue:
 - a. Si se está en la primera fila de la tabla de datos asignar al contador de enfermedades el valor de 1, asignar la variable de enfermedad comparadora mediante un Get Row Item y finalmente asignar a la tabla de datos en el campo “conteo” el valor del contador de enfermedades con la actividad Assign.
 - b. Si la condición previa prueba falsa entonces comparar si la enfermedad actual es igual a la enfermedad comparadora, si lo es

incrementar el contador de enfermedades en una unidad y asignarla a la tabla de resultados.

- c. Si la enfermedad no es igual a la anterior entonces se procede a asignar el acumulador de enfermedades total mediante un Assign. Se asigna el porcentaje como el total de conteo de enfermedades actuales dividiéndolo por la longitud de la base de datos. Se asigna el campo "Porcentaje" en la fila anterior, donde la enfermedad todavía no había cambiado. Se continúa asignando de nuevo Contadores, el campo "Conteo" y finalmente la enfermedad comparadora, mediante actividades de Assign y Get Row Item respectivamente.

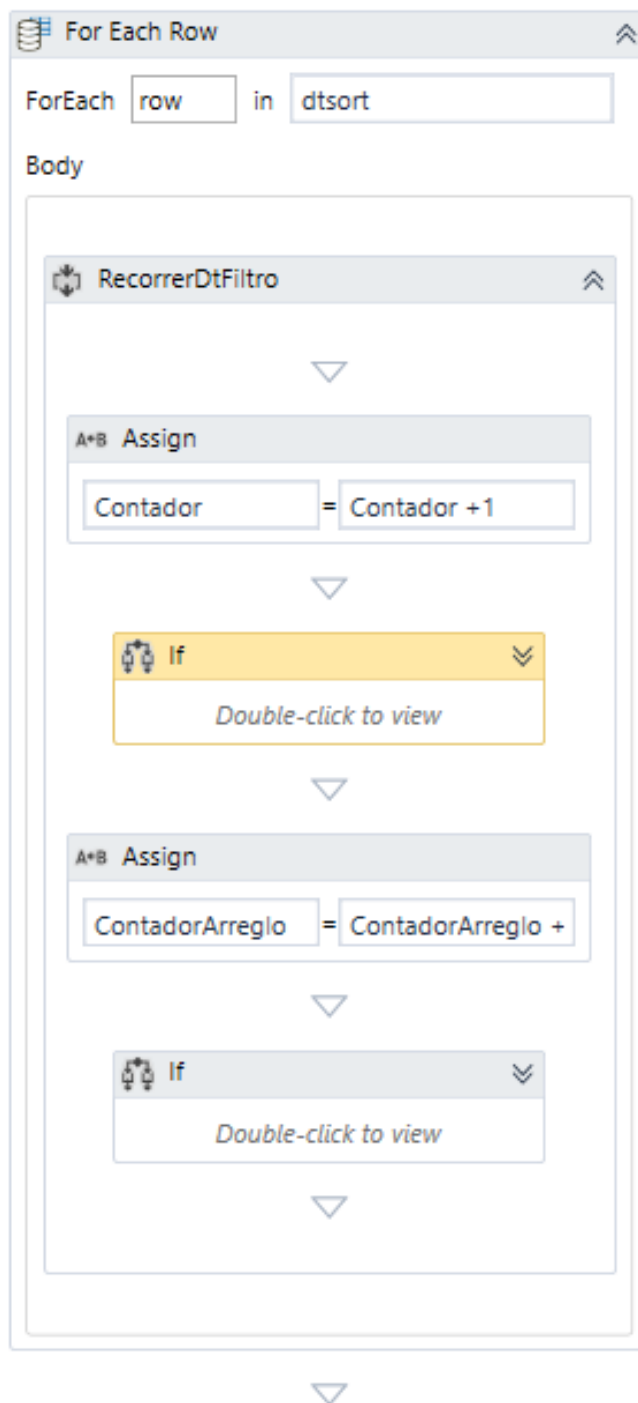


Figura 9: For Each Row descrito en la actividad 5. Fuente: Autoría propia

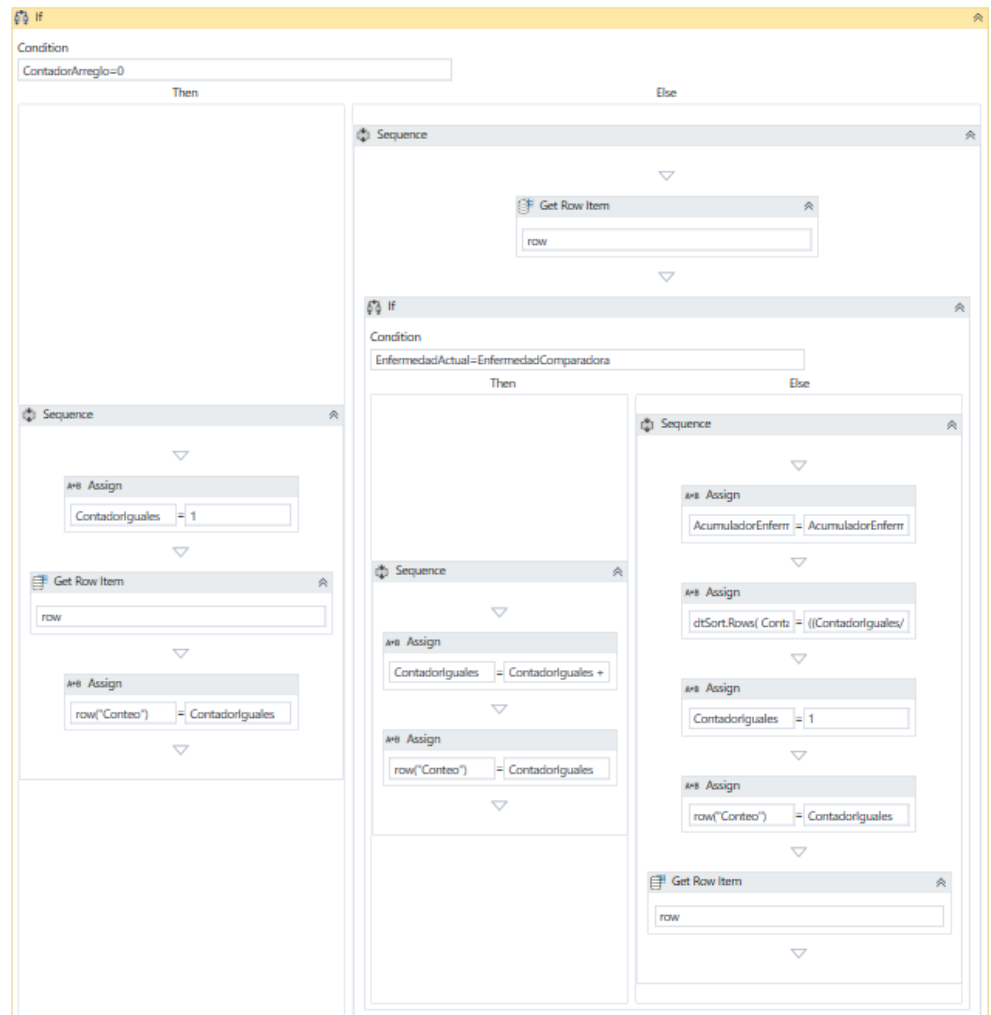


Figura 10: Muestra los condicionales descritos en las actividades 5a, 5b y 5c. Fuente: Autoría propia

6. Todavía dentro del for each row previo se acumula la variable ContadorArreglo mediante una actividad assign.
7. Se aplica nuevamente otra condición en donde sí se está en la última fila y la enfermedad actual es igual a la comparadora entonces se asigna el conteo y el porcentaje correspondiente a la tabla. Si esta condición no se cumple entonces sigue otra condición lógica en la cual se pregunta únicamente si el índice de recorrido de la tabla de datos es la última. Si sí, significa que la enfermedad comparadora es distinta de la enfermedad actual, por lo que se asigna el porcentaje respectivo en la fila actual a la tabla y se acumula la variable acumuladora mediante un Assign.

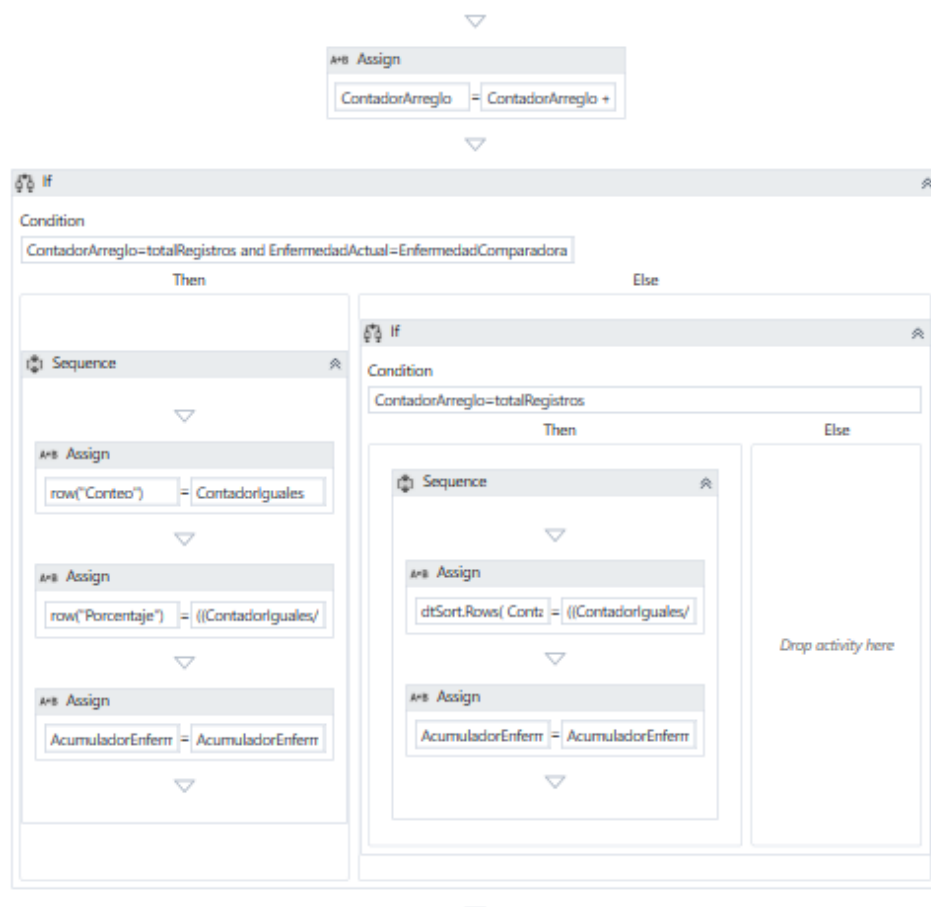


Figura 11: Actividades 6 y 7. Fuente: Autoría propia

8. Después de la actividad del For each Row en donde se adquieren las frecuencias de las enfermedades, se procede a filtrar los porcentajes vacíos mediante la actividad Filter data table, dejando únicamente los valores con porcentaje.
9. Se utiliza una actividad Sort Data table donde se ordena la tabla de resultados de manera descendente por la columna "Porcentaje".
10. Se continúa con la actividad Build Data Table en donde se genera una tabla de resumen de morbilidad. Donde posteriormente se asignan los primeros dos diagnósticos más frecuentes mediante las actividades Add Data Row.
11. Se añade una fila a la tabla de validación mediante la actividad Add Data Row con los valores de la longitud de la tabla y los valores acumulados de enfermedades.
12. Finalmente se utilizan tres actividades de Write Range, en donde se escriben las tres tablas de datos; la primera del resultado de morbilidad general, la segunda sobre el resumen de los dos diagnósticos más frecuentes y finalmente una tabla de validación, la cual contiene el total de registros de la tabla con el conteo de enfermedades acumulado. Generando los archivos

de Excel; Resultado_MorbilidadGeneral, ResumenMorbiladGeneral y finalmente ValidacionGeneral.

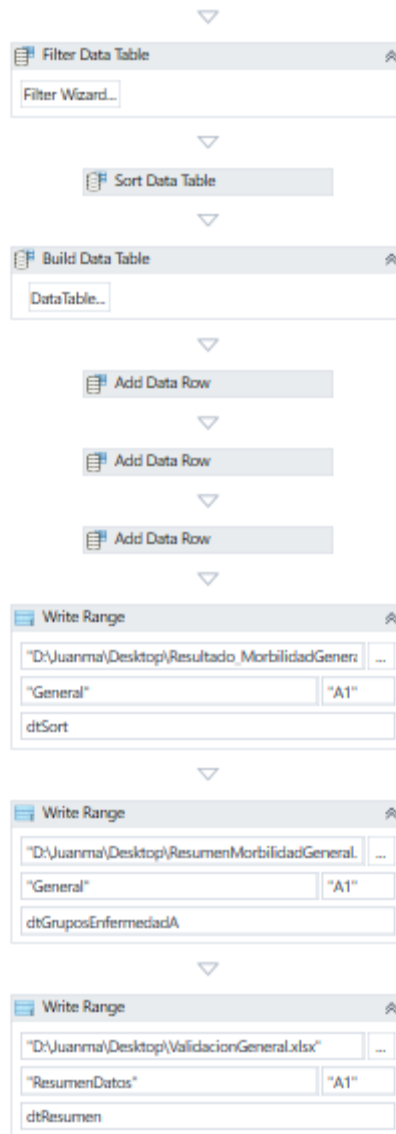


Figura 12: Actividades 8-12. Fuente: Autoría propia

Etapas 2. Filtrar Enfermedades

La segunda etapa de la Segmentación General de la población cuenta con una secuencia llamada Filtrar Enfermedades. Esta consiste en realizar sobre el archivo de Excel Datos, filtros con las enfermedades más relevantes según el archivo de Resumen de morbilidad general y generar otros archivos para futuro tratamiento con estos filtros.

Las actividades de Filtrar Enfermedades son las siguientes:

1. Leer el archivo de Excel Datos mediante la actividad Read Range. Este es un archivo que contiene otras columnas de interés fuera de las de Datos RIPS, el cual tiene una estructura predeterminada para realizar la caracterización de la morbilidad.
2. Leer el archivo de ResumenMorbilidadGeneral mediante una actividad Read Range.
3. Un for each row con sobre la tabla de datos del literal 2 con las siguientes actividades:
 - a. Get row item en donde se extrae la enfermedad a filtrar mediante un Filter Data table.
 - b. Un filtro mediante la actividad Filter Data Table en donde se filtra la enfermedad adquirida.
 - c. Una actividad de Assign en donde se asigna los primeros cuatro códigos de la clasificación IDC para escribir el archivo.
 - d. Finalmente una actividad de Write range en donde se escribe a una carpeta llamada ABCySeriesTiempo el archivo con nombre "Filtro" más el código extraído de la enfermedad.

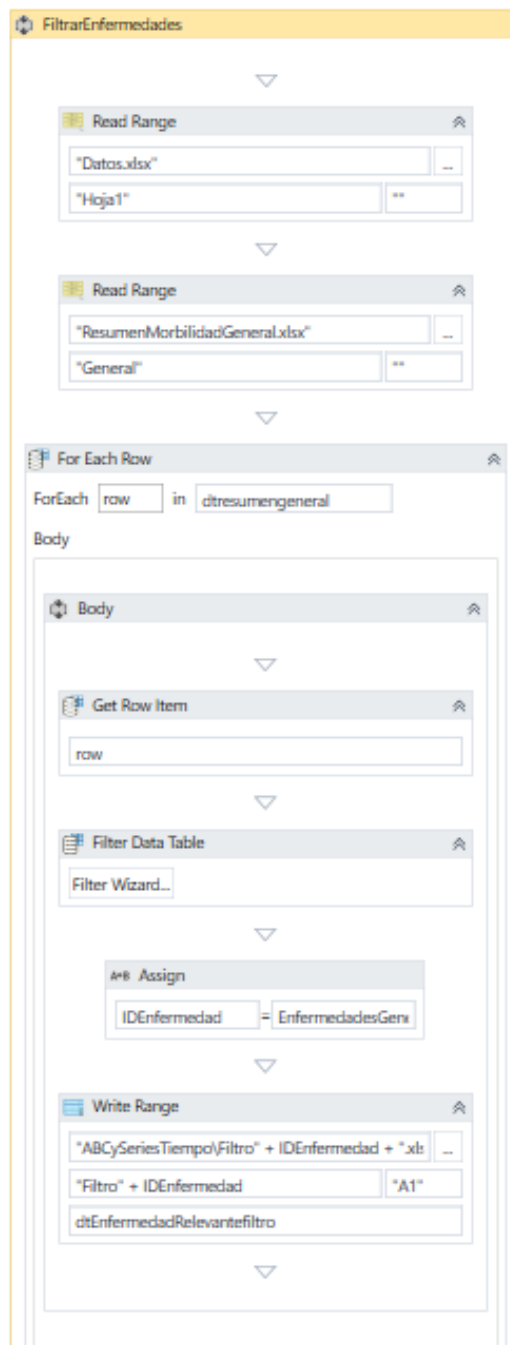


Figura 13: Etapa 2 FiltrarEnfermedades. Fuente: Autoría propia

Etapa3: Generación Datos

Las etapas 1 y 2 consistieron totalmente en actividades de UiPath utilizando fundamentos de lógica y conocimiento sobre las actividades. La Etapa 3 llamada "GeneraciónDatos" utiliza como herramienta principal las actividades de UiPath para

introducir SQL a los archivos de Excel. Esta etapa genera dentro de los archivos de los filtros de las enfermedades otras 4 hojas de Excel. Estas siendo: segmentación ABC de procedimientos, comportamiento en el tiempo de diagnósticos con finalidad y costo y finalmente un conteo de ocurrencias del diagnóstico filtrado por municipio de residencia.

Las actividades utilizadas para la etapa de Generación de datos son las siguientes:

1. Una actividad Read Range para leer la tabla de datos del Resumen de Morbilidad General.

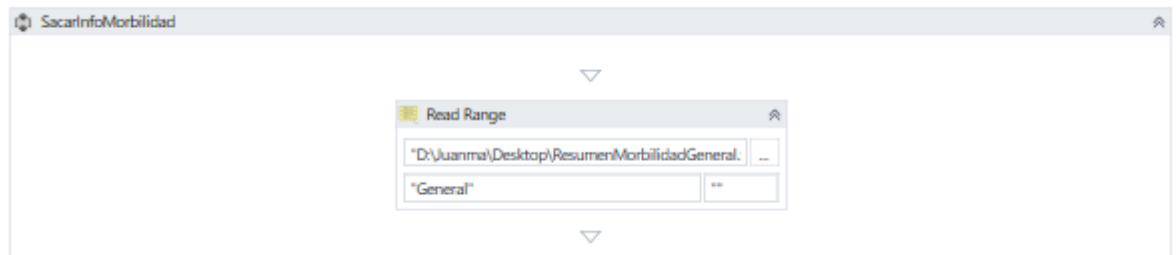


Figura 14: Actividad 1. Fuente: Autoría propia

2. Una actividad For Each Row donde se recorre la tabla de datos del literal 1. Al ser un estudio para varios diagnósticos la generación de conteos y agrupaciones se hace en ciclo para las enfermedades del resumen de morbilidad. A continuación, se describen las actividades dentro del ciclo For Each Row.
 - a. Se define un contador con la actividad assign, el cual se va acumulando con cada iteración del ciclo.
 - b. Se extrae el diagnóstico al cual se le van a segmentar y agrupar datos mediante un Get Row Item en la Columna nombrada "Enfermedad" de la tabla en cuestión.
 - c. Un Assign en donde se extraen los primeros 4 caracteres del código del diagnóstico adquirido en el literal 2, llamándolo CodigoEnfermedad.
 - d. Se procede a definir mediante un Assign una cadena que contiene la conexión a la hoja de Excel donde se encuentra el filtro. Esta cadena toma el valor "D:***** \Desktop\ABCySeriesTiempo\Filtro" + CodigoEnfermedad + ".xlsx"

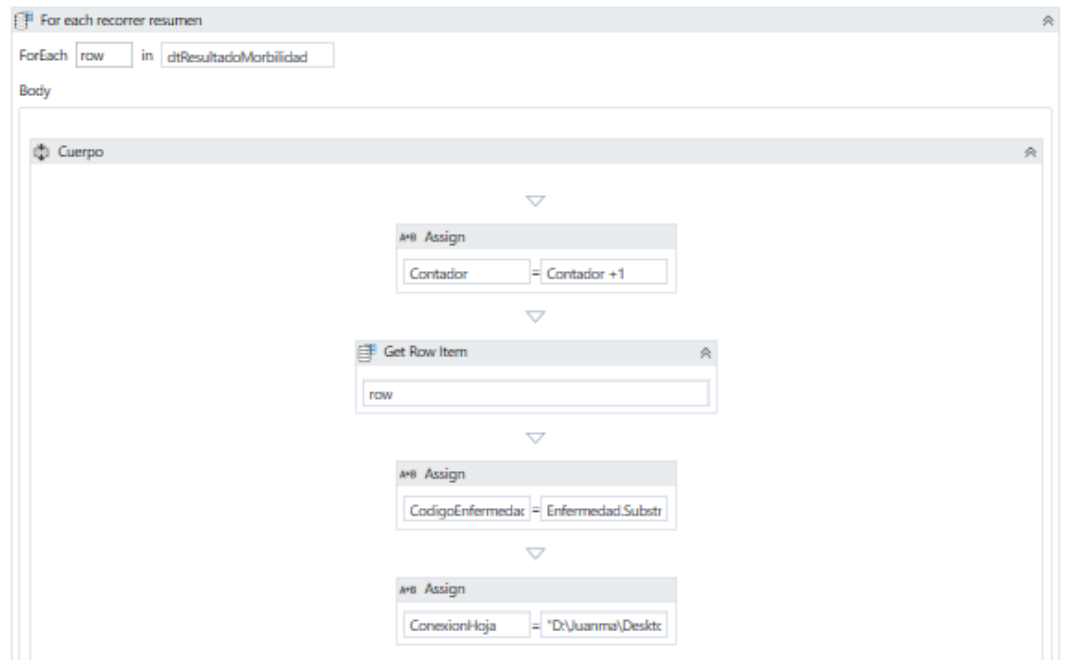


Figura 15: Actividades a, b, c y d. Fuente: Autoría propia

- e. Al realizar la conexión sigue una secuencia que contienen las distintas consultas a realizar mediante SQL. Los códigos de cada consulta se hacen con el lenguaje de SQL Access y son como sigue:
- i. "SELECT ([ID Procedimiento]), COUNT([ID Procedimiento]) AS ConteoProcedimientos, SUM([Costo Procedimiento]) AS CostoProcedimientos FROM [Filtro" + CodigoEnfermedad + "\$]T1 WHERE [Costo Procedimiento] >0 GROUP BY ([ID Procedimiento]) ; "
 Este es un código que busca seleccionar todos los procedimientos con costo mayor a 0, es decir que el tipo de atención sea un procedimiento de salud con costo mayor a 0, haga un conteo por cada procedimiento y sume los costos totales de los procedimientos. Esta consulta genera una tabla de datos llamada dtABC.
 - ii. "SELECT ([ID Procedimiento]), [Mes], [Finalidad Procedimiento], COUNT([ID Procedimiento])AS ConteoProcedimientos, SUM([Costo Procedimiento])AS CostoProcedimientos FROM [Filtro" + CodigoEnfermedad + "\$] T1 WHERE ([Costo Procedimiento]) > 0 GROUP BY ([ID Procedimiento]), Mes, [Finalidad Procedimiento] ORDER BY ([ID Procedimiento]), Mes "

Este es un código de consulta que busca seleccionar el código del procedimiento, mes, finalidad del procedimiento, la cuenta de procedimientos y finalmente la suma del código del procedimiento, todo esto agrupado por procedimiento, mes, y finalidad. Esto con el fin de poder analizar el comportamiento en el tiempo del procedimiento con su respectiva finalidad, costo y cantidad. Esta consulta genera una tabla de datos llamada dtSerieTiempoconCosto.

- iii. "TRANSFORM COUNT(Enfermedad) SELECT (Enfermedad) FROM [Filtro" + CodigoEnfermedad + "\$] T1 GROUP BY (Enfermedad) PIVOT T1.Mes"

Esta consulta busca contar la ocurrencia de cada diagnóstico del filtro actual y transponer este conteo por mes. Con el fin de generar el comportamiento en el tiempo de la enfermedad. El resultado de dicha consulta es una tabla de datos llamada dtSerieTiempoEnfermedades.

- iv. " SELECT ([Municipio Residencia]), COUNT([Municipio Residencia]) AS ConteoMunicipio FROM [Filtro" + CodigoEnfermedad + "\$] T1 GROUP BY ([Municipio Residencia]) ORDER BY COUNT([Municipio Residencia]) DESC"

La última consulta selecciona el Municipio de residencia, cuenta las veces que se presentó este municipio de residencia en el filtro del diagnóstico como ConteoMunicipio, agrupa este conteo por municipio y finalmente ordena este conteo de manera descendente. Esto con el fin de identificar cuáles son los municipios donde más se presenta esta enfermedad. La tabla de datos generada por esta consulta se llama dtmunicipios.

- f. Después de generar las consultas de interés mediante los queries se procede a desconectar la conexión creada a la hoja de Excel donde mediante una actividad Disconnect.

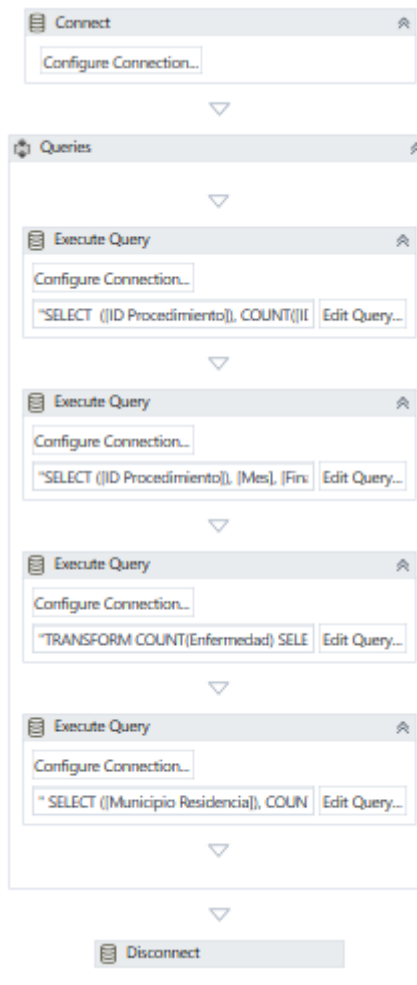


Figura 16: Actividades e y f. Fuente: Autoría propia

- g. Seguido a esta actividad se encuentra una secuencia en donde se genera el resultado final de la tabla ABC. Por simplicidad este se metió dentro de una secuencia, pero las actividades dentro de este son:
 - i. En primera instancia se encuentra una actividad Sort Data Table, en donde se orden los costos de los procedimientos.
 - ii. Luego se encuentran 5 actividades de Add Data Column, en donde se añaden las 5 de tipo doble. Estas columnas vienen siendo "PorcentajeIDProcedimientos", "PorcentajeCostoTotal", "PorcentajeIDAcumulados", "PorcentajeCostoAcumulado" y finalmente "Clasificación".
 - iii. A estas actividades las sigue una actividad de assign, en donde se asigna la variable LongBD, la cual representa la longitud de la tabla de datos de la tabla de datos ABC.

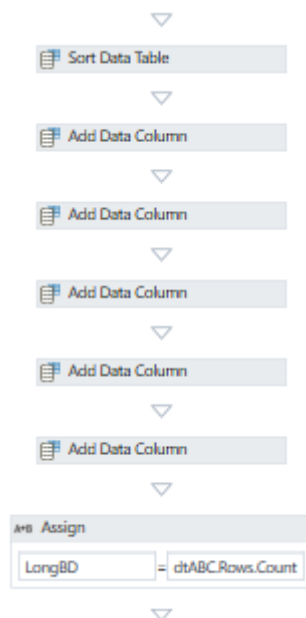


Figura 17: Actividades g.i, g.ii y g.iii. Fuente: Autoría propia

- iv. Dentro de una actividad For Each Row independiente se tiene una variable llamada CostoAcumuladoProcedimiento la cual es una variable acumuladora que va guardando cada valor del costo y sumándolo a un costo total.

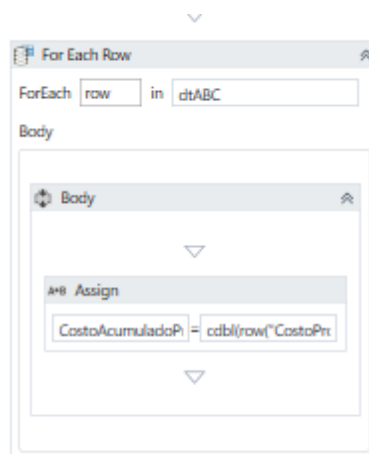


Figura 18: Actividad g.iv. Fuente: Autoría propia

- v. Luego de haber conseguido este costo total se procede a una actividad For Each Row. En esta se define un contador llamado

ContadorFilas. Posteriormente se asigna mediante unas actividades de assign los valores a las siguientes columnas:

1. Row("PorcentajeIDProcedimiento")=Cdbl(1/LongBD)
2. Row("PorcentajeCosto") = cdbl(row("CostoProcedimientos") / CostoAcumuladoProcedimientos)*100
3. CostoPorcentajeAcumulado = CostoPorcentajeAcumulado + cdbl(row("PorcentajeCosto").tostring)
4. AcumuladoOcurriencia = cdbl(ContadorFilas/cdbl(LongBD))
5. row("PorcentajeCostoAcumulado") = CostoPorcentajeAcumulado
6. row("PorcentajeIDAcumulado") = AcumuladoOcurriencia*100

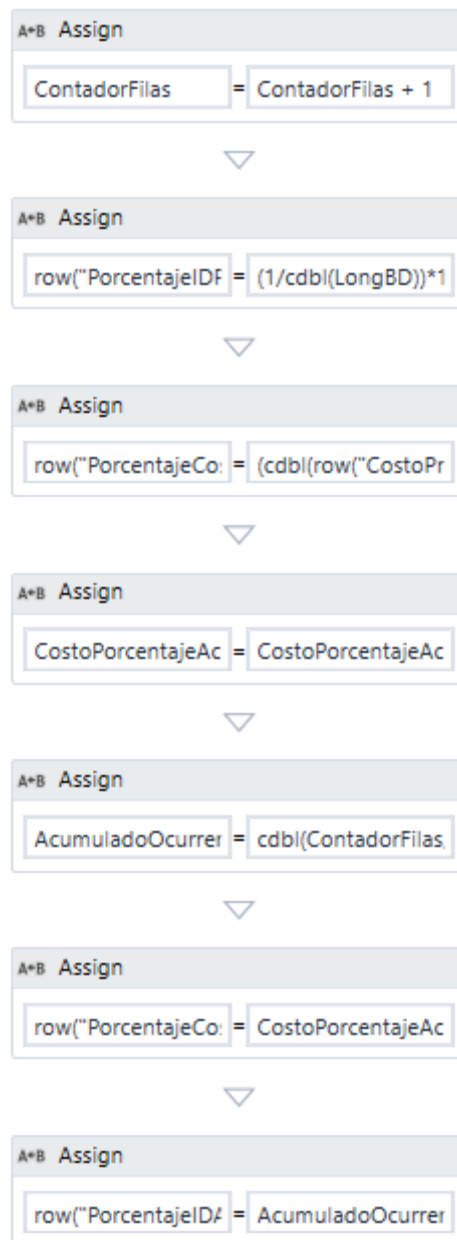


Figura 19: Actividades g.v. Fuente: Autoría propia

vi. Por último, se tiene dentro del For Each Row una actividad condicional IF. En esta se tiene la siguiente condición:

`cdbl(row("PorcentajeIDAcumulado").ToString) < 20`

Es decir, si el porcentaje acumulado de ID's de enfermedades es menor al 20% estas enfermedades representan la Clasificación A. Si esta condición prueba verdadera se asigna el valor correspondiente dentro de la secuencia del THEN como sigue:

```
row("Clasificación") = "A"
```

Por otro lado, si esta condición es falsa, la porción del Else tiene otro condicional IF, el cual tiene la siguiente condición lógica:

```
cdbl(row("PorcentajeIDAcumulado").ToString) < 50
```

Esta condición verifica si el porcentaje de ID de procedimientos acumulados es menor al 50%.

Si en esta fila se cumple esta condición se asigna el siguiente valor en la porción del Else mediante una actividad de Assign:

```
row("Clasificación") = "B"
```

Si la condición no se cumple en la porción del Else se asigna el siguiente valor:

```
row("Clasificación") = "C"
```

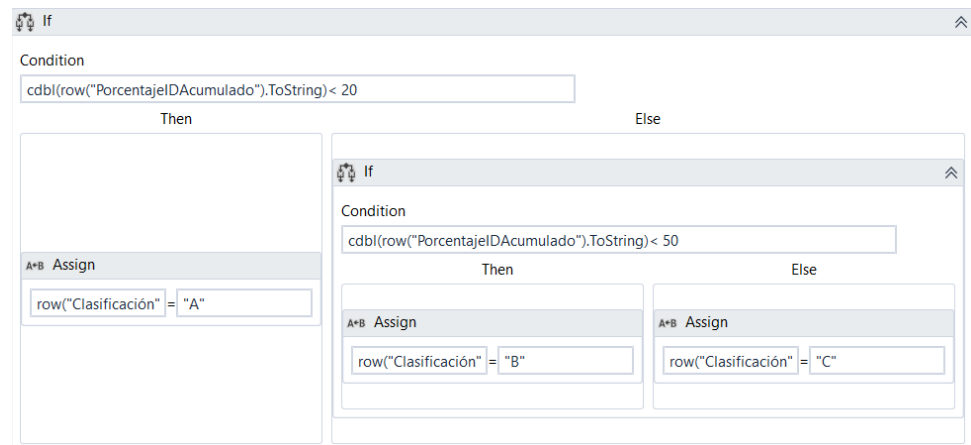


Figura 20: Actividad g.vi. Fuente: Autoría propia

- h. Finalmente se tiene la secuencia EscribirResultados en la cual se tienen cuatro actividades de Write Range. Para el caso del estudio general estas todas tienen como ruta final la hoja de Excel:

"D:\Juanma\Desktop\ABCySeriesTiempo\Filtro" + CodigoEnfermedad
+".xlsx"

Siendo CodigoEnfermedad la misma variable definida para escribir el filtro.

Se escriben las siguientes hojas con las siguientes tablas de datos;

"ABC" con dtABC

"STProcedimientosCosto" con dtSerieTiempoConCosto

"STEnfermedad" con dtSerieTiempoEnfermedades

"ConteoMunicipios" con dtMunicipios

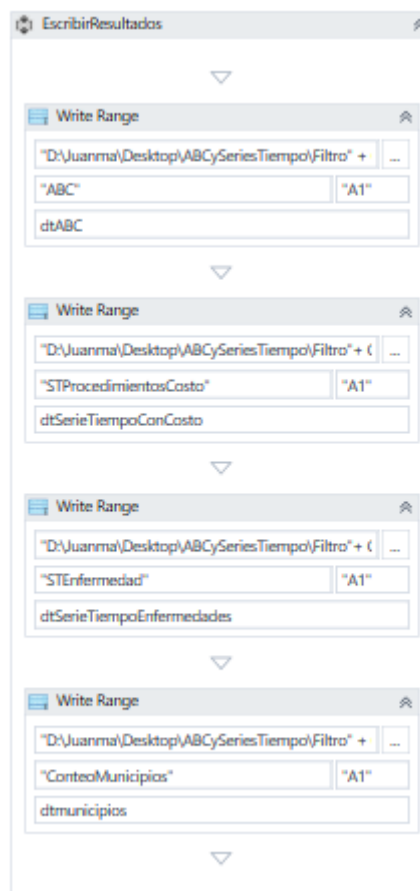


Figura 21: Actividad h. Fuente: Autoría propia

Con estas actividades de Write range dentro del For each Row ya se finaliza la etapa de GeneraciónDatos, con estos archivos ya se puede

proceder a hacer un análisis y posteriormente más tratamiento de acuerdo con la información que se desee extraer del diagnóstico en cuestión.

Desarrollo RPA:

Segmentación por Grupos

La segmentación por grupos cuenta con las mismas etapas que tiene la segmentación general, solo que el tratamiento de datos está orientado a los grupos, entonces los pasos se hacen dentro de un ciclo para cada uno.

Las diferencias para cada etapa son como siguen:

Etaa 1 Caracterización Morbilidad Por grupos:

En comparación a la caracterización general, a la caracterización por grupos en las actividades tiene las siguientes actividades después de la actividad 2 que generan una modificación al desarrollo:

1. Una actividad read range en donde se lee el archivo con la información por grupos, siendo este el archivo Prueba_DatosRips.xlsx. Esta actividad genera una tabla de datos llamada dtgrupos.

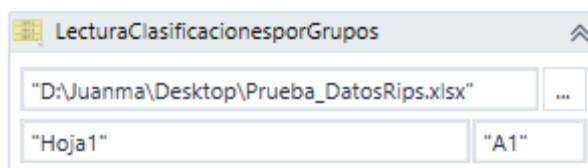


Figura 22: Actividad lectura Prueba_DatosRips.xlsx. Fuente: Autoría propia

2. Luego se encapsulan todas las actividades desde la 3 hasta la 11 de la Caracterización general dentro de una actividad For Each Row dentro de la tabla Grupos.

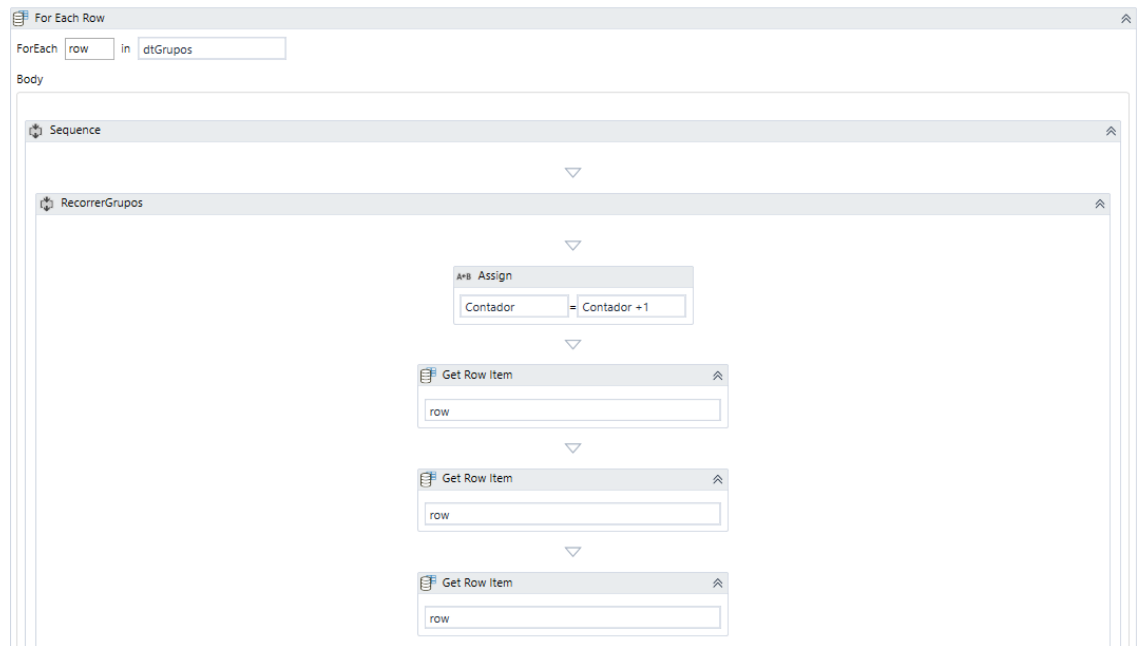


Figura 23: Literales 2 y 3. Fuente: Autoría propia

3. Dentro de este for each row se vuelve a asignar un contador y se traen 4 elementos necesarios de los grupos de etapas de ciclo vital. Estos cuatro se traen mediante las actividades Get Row Item con los valores de las columnas y variables como siguen:
 - “Grupo” como grupo
 - “EdadInf” como Edad
 - “EdadSup” como Edad1
 - “Genero” como Genero
4. Con estas variables asignadas se procede a hacer una actividad de Filter Data Table sobre la tabla de datos llamada dtDatos. Siendo esta la misma tabla leída en el literal 1 de la caracterización general, actividad que se tiene tanto para el estudio general como el descrito. En esta actividad se filtran todos los valores del grupo en donde la columna “Edad”>=Edad, “Edad”<=Edad y finalmente “Genero”=Genero. Esta tabla después procede a ser ordenada y sobre esta se realiza la caracterización de la morbilidad para cada grupo, es decir el procedimiento donde se hace un recuento de diagnósticos, se les asigna un porcentaje y finalmente se ordenan de manera descendente para obtener las enfermedades más frecuentes.

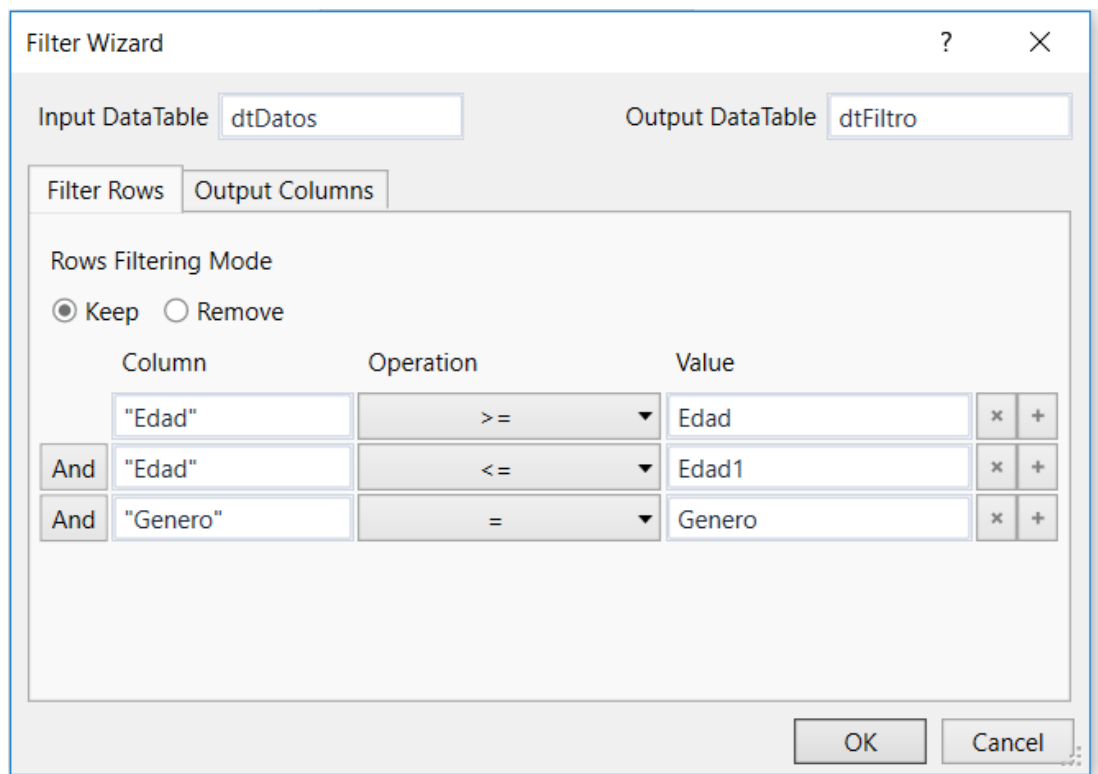


Figura 24: Configuración de la actividad Filter Data Table. Fuente: Autoría propia

5. Posterior a este filtro las actividades 2 hasta la 11 de la caracterización general se repiten exactamente igual, la única diferencia es que ahora se hacen dentro de un ciclo y la tabla dtfiltro, resultado del filtro se ordena y sobre esta se hace el tratamiento de datos.
6. La actividad 12 sufre una pequeña variación, ya que los 3 archivos de resultado cambian de nombre. Estos reciben los nombres de Resultado_Morbilidad, ResumenMorbilidad y ResumenValidación. Además de esto la actividad Write Range del Resumen Validación se encuentra fuera de la actividad For Each Row de la tabla grupos, con el fin de evitar que el archivo se escriba múltiples veces sin necesidad.

Etapa 2 Filtro por Grupos y Enfermedades:

Al igual que la etapa 1 para la segmentación por grupos, la etapa 2 es muy similar a la caracterización general. De nuevo se generan los mismos resultados sobre los grupos basándose en el ResumenMorbilidad. Es decir, se repite la misma estructura en la etapa 1 que en la dos. Se lee el archivo de Datos.xlsx, el archivo de Prueba_DatosRips.xlsx y se generan los filtros por grupo y por enfermedad. En esta situación las diferencias a recalcar son las siguientes.

1. Después de la lectura del archivo de Prueba_DatosRips, se tienen una actividad For Each Row para los grupos.
2. Dentro de esta actividad se encuentra un contador, el cual indica en qué posición de la tabla de datos se está, es decir el número de grupos. Este incrementa mediante una actividad Assign.
3. Seguido al assign, dentro del For Each Row de los grupos se tiene una actividad IF. La cual busca comprobar si el contador es mayor a 6 y menor a 9. Esta condición se hace con el fin de limitar los resultados del RPA. Si es verdadera significa que se tienen en específico los grupos 7 y 8. Grupos que posteriormente van estarán sujetos a análisis.
4. Si la condición se cumple se tiene una actividad de Read Range el cual lee la hoja de ResumenMorbilidad. Este archivo sirve de base para seleccionar las dos enfermedades más comunes de cada grupo al igual como ocurrió para el filtro en la segmentación general.
5. Luego se adquieren mediante 5 actividades de Get Row Item las variables de Grupo, Edad, Edad1, Genero y Enfermedad de la tabla de ResumenMorbilidad.
6. Se procede a utilizar un Filter Data Table para filtrar el grupo y diagnóstico en cuestión de los datos originales.
7. Finalmente, la actividad Write Range escribe a las hojas de Excel dentro de la misma carpeta, pero esta vez con nombre "Grupo" + Grupo + CodigoEnfermedad y con nombre de hoja como "Filtro" + CodigoEnfermedad.

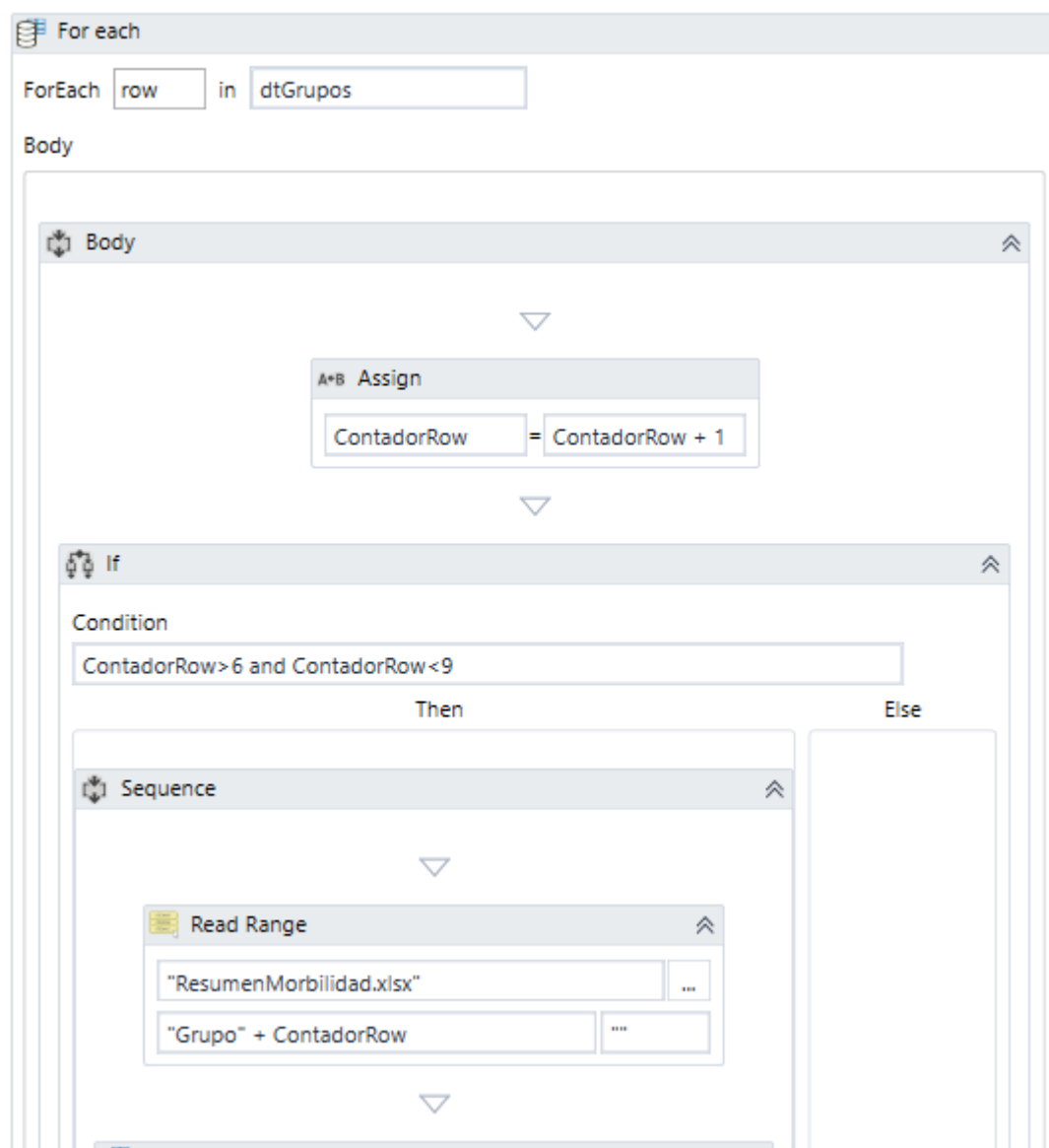


Figura 25: Literales 1 y 2. Fuente: Autoría propia

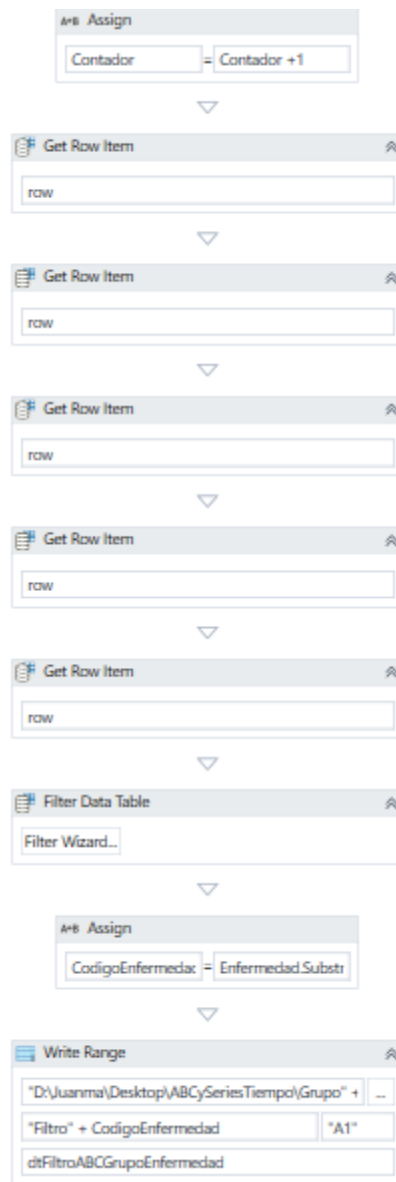


Figura 26: Literales 3-7. Fuente: Autoría propia

Etapa 3 Generación Datos por Grupo:

La etapa 3 al igual que la etapa 1 y 2, sigue un proceso iterativo, en donde se lee de nuevo la hoja Prueba_DatosRIPS y se extrae la información relevante de los grupos. Con base en esto se crea la conexión a la hoja de Excel mediante SQL y se repite el mismo tratamiento de datos que la generación general solo que esta vez para cada grupo. En este caso es pertinente hacer la lectura del resumen ya que para conectarse a

los archivos de Excel de manera iterativa es necesario contar con el nombre de la hoja y el nombre del archivo.

Análisis en Series de Tiempo:

Una vez se obtienen estos resúmenes en Excel, se tienen los datos necesarios para terminar su análisis.

Estos datos son llevados a R para realizar las series de tiempo. En R, se utilizan las bibliotecas TSA y Cairo para descomponer las series de tiempo en tendencia y estacionalidad. Además, se realizan las gráficas para la inspección visual de dicha descomposición.

Se utiliza el código en el Anexo 2 para la ejecución del análisis.

Desarrollo VaR:

Para calcular el capital requerido mediante simulación Montecarlo se utilizó la aplicación Risk Palisade. Se definió un nivel de $\alpha = 0,1\%$ tomando el valor recomendado de (Flores, 2012). Este método de cálculo del capital requerido se aplicó sobre los procedimientos 9073 y 12211 con el fin de determinar el capital a tener en caso de un evento extremo. Como entrada de frecuencia se tomaron los datos de demanda mensual del procedimiento mientras que la entrada de severidad se determinó como el costo de cada uno de estos procedimientos.

En este caso específico, asumiendo independencia dada la naturaleza de los procedimientos, al tener la frecuencia mensual y la severidad por procedimiento, para la simulación fue necesario simular la severidad para cada uno de los procedimientos dada la naturaleza de los datos, con el fin de que cada procedimiento estuviera sujeto a su respectiva distribución de severidad. Al fin y al cabo, se sumaron estos costos para obtener el LPA.

En primera instancia con el resultado del RPA en la hoja de Excel STProcedimientoCostoFinalidad se procedió a elegir la demanda mensual de los procedimientos para el 2017. Los datos son como se muestran en las siguientes tablas:

ID Procedimiento	Mes	Finalidad Procedimiento	ConteoProcedimientos
9073	1/01/2017	2 - TERAPÉUTICO	8
9073	1/02/2017	2 - TERAPÉUTICO	29
9073	1/03/2017	2 - TERAPÉUTICO	19
9073	1/04/2017	2 - TERAPÉUTICO	22
9073	1/05/2017	2 - TERAPÉUTICO	14
9073	1/06/2017	2 - TERAPÉUTICO	12
9073	1/07/2017	2 - TERAPÉUTICO	10
9073	1/08/2017	2 - TERAPÉUTICO	11
9073	1/09/2017	2 - TERAPÉUTICO	8
9073	1/10/2017	2 - TERAPÉUTICO	5
9073	1/11/2017	2 - TERAPÉUTICO	8
9073	1/12/2017	2 - TERAPÉUTICO	3

Tabla 6: Demanda mensual para el ID procedimiento 9073 con finalidad Terapéutico para el Grupo 8 con diagnóstico O800 en el año 2017. Obtenido de: autoría propia basada en datos RIPS

ID Procedimiento	Mes	Finalidad Procedimiento	ConteoProcedimientos
12211	1/01/2017	1 - DIAGNÓSTICO	6
12211	1/02/2017	1 - DIAGNÓSTICO	5
12211	1/03/2017	1 - DIAGNÓSTICO	4
12211	1/04/2017	1 - DIAGNÓSTICO	9
12211	1/05/2017	1 - DIAGNÓSTICO	26
12211	1/06/2017	1 - DIAGNÓSTICO	19
12211	1/07/2017	1 - DIAGNÓSTICO	7
12211	1/08/2017	1 - DIAGNÓSTICO	0
12211	1/09/2017	1 - DIAGNÓSTICO	0
12211	1/10/2017	1 - DIAGNÓSTICO	0
12211	1/11/2017	1 - DIAGNÓSTICO	0
12211	1/12/2017	1 - DIAGNÓSTICO	0

Tabla 7: Demanda mensual para el ID procedimiento 12211 con finalidad Diagnóstico para el diagnóstico Z139 para la población general. Obtenido de: autoría propia basada en datos RIPS

Luego se procedió a extraer los costos del procedimiento en cuestión del filtro indicado para el año 2017. Esto se hizo tanto para la población general con diagnóstico Z139 como para el diagnóstico O800 del Grupo 8 F. Estos datos se encuentran en el anexo 5.

El procedimiento 9073 en cuestión se presentó 149 veces para el año 2017 y se adquirieron respectivamente 149 datos de costos de los procedimientos. Por otro lado, el procedimiento 12211 analizado se presentó 76 veces con la misma cantidad de entradas para el costo.

Luego se procedió a realizar un análisis de la estadística descriptiva, histogramas y pruebas de bondad de ajuste para los datos de frecuencia y costo de los procedimientos. Esto se hizo mediante la aplicación Statfit, la cual es una herramienta de análisis estadístico que permite ajustar datos, realiza análisis descriptivos y arroja resultados gráficos sobre una muestra de datos ingresada.

Para esta situación las pruebas de bondad de ajuste se realizaron con estimadores de máxima verosimilitud y con pruebas Anderson-Darling y Kolmogorov-Smirnov.

A continuación, se demuestran las estadísticas descriptivas, histogramas y ajuste para el procedimiento 9073:

descriptive statistics	
data points	12
minimum	3.
maximum	29.
mean	12.4167
median	10.5
mode	16.
standard deviation	7.52521
variance	56.6288
coefficient of variation	60.6057
skewness	1.07812
kurtosis	0.15581

Figura 27: Estadística descriptiva demanda mensual (Frecuencia) ID 9073 procedimiento con finalidad terapéutica para el año 2017. Obtenido de: autoría propia

descriptive statistics	
data points	149
minimum	434358
maximum	1.2762e+006
mean	589885
median	542385
mode	542385
standard deviation	143015
variance	2.04532e+010
coefficient of variation	24.2445
skewness	3.26788
kurtosis	9.54693

Figura 28: Estadística descriptiva para la severidad del 2017 del procedimiento 9073 con finalidad terapéutica. Obtenido de: autoría propia

Con la naturaleza del comportamiento de la frecuencia de las visitas se esperaba que la distribución de esta fuera una Poisson. Analizando la estadística descriptiva se buscaba que la media fuera similar a la varianza, pero esto no sucedió. Más sin embargo el comportamiento no se descartó y se procedió a hacer una prueba de bondad de ajuste.

Por otro lado, se esperaba según la teoría que la severidad siguiese una distribución continua, como es de esperarse en los costos de un procedimiento. Pero como se puede observar con los datos de sesgo y de curtosis el comportamiento del costo para este procedimiento no apunta una distribución ni normal ni log normal y aparenta ser una distribución discreta dados los valores fijos que toma.

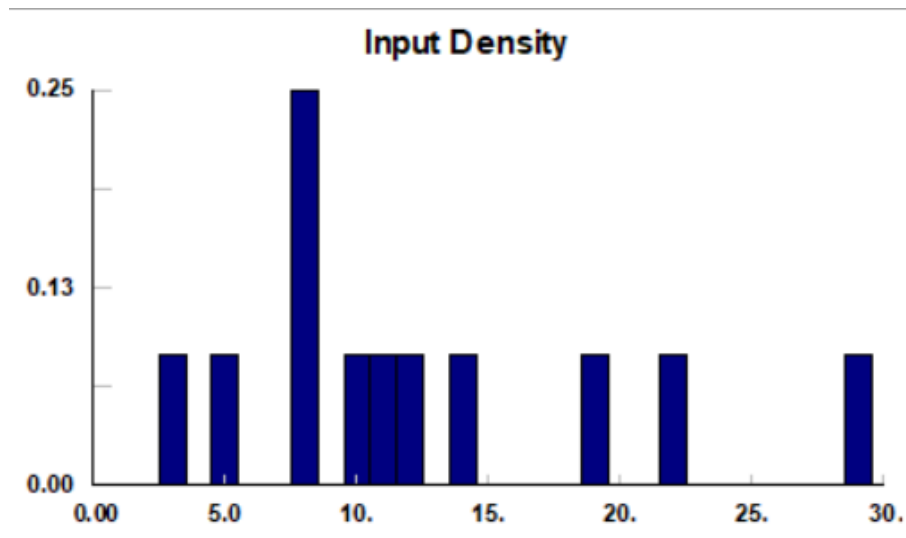


Figura 29: Histograma para el comportamiento de la demanda del procedimiento 9073. Obtenido de: autoría propia

Según el histograma, con un análisis visual sencillo se puede proceder a determinar que el comportamiento de los datos se asemeja a una Poisson con datos más frecuentes cerca de 10 visitas mensuales, corroborando la media de 12.42 de la estadística descriptiva. Además, se ve una concentración de datos cerca de la media y una cola hacia la derecha indicando también el comportamiento de una distribución de Poisson.

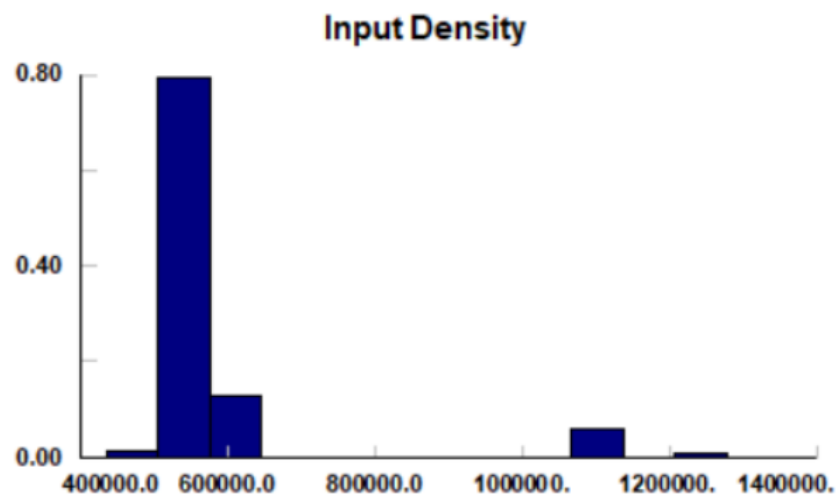


Figura 30: Histograma de los costos para el procedimiento 9073 analizado. Obtenido de: autoría propia

Por otro lado, el histograma de los costos indica más bien que los datos se ajustan a una distribución discreta empírica, esto se debe a que se encuentran una cantidad finita de costos con una frecuencia variable, indicando que no se tienen infinitos valores de costo sino más bien unos cuantos valores discretos que puede tomar la variable. Esto se puede identificar fácilmente con los valores extremos donde el costo del procedimiento es superior a 1.000.0000. Estos valores extremos y la coincidencia de clases concretas indican la presencia de una distribución discreta.

Auto::Fit of Distributions		
distribution	rank	acceptance
Poisson[12.4]	100	do not reject
Geometric[7.45e-002]	88.6	do not reject
Discrete Uniform[3., 29.]	75.5	do not reject

Figura 31: Pruebas de bondad de ajuste demanda procedimiento 9073 analizado. Obtenido de: autoría propia

Como se muestra en las pruebas de bondad de ajuste estas indican que no se puede rechazar el comportamiento de la demanda como una distribución Poisson con media 12,42.

Auto::Fit of Distributions		
distribution	rank	acceptance
Beta(4.34e+005, 1.28e+006, 3.58e-002, 0.451)	0.	reject
Erlang(4.34e+005, 3., 4.85e+004)	0.	reject
Exponential(4.34e+005, 1.56e+005)	0.	reject
Gamma(4.34e+005, 3.21, 4.85e+004)	0.	reject
Inverse Gaussian(4.34e+005, 5.66e+005, 1.56e	0.	reject
LogLogistic(4.34e+005, 2.35, 1.25e+005)	0.	reject
Lognormal(4.34e+005, 11.8, 0.482)	0.	reject
Pareto(4.34e+005, 3.5)	0.	reject
Pearson 5(4.34e+005, 6.49, 7.92e+005)	0.	reject
Pearson 6(4.34e+005, 3.85e+005, 2.12, 5.94)	0.	reject
Triangular(4.34e+005, 1.28e+006, 4.34e+005)	0.	reject
Uniform(4.34e+005, 1.28e+006)	0.	reject
Weibull(4.34e+005, 1.41, 1.76e+005)	0.	reject
Johnson SB	no fit	reject
Rayleigh(4.34e+005, 1.5e+005)	0.	reject
Inverse Weibull(4.34e+005, 5.11, 8.88e-006)	0.	reject
Chi Squared	no fit	reject
Power Function(4.34e+005, 1.36e+006, 0.515)	0.	reject

Figura 32: Prueba de bondad de ajuste Severidad 9073. Obtenido de: autoría propia

Como se puede observar la severidad no se ajustó a ninguna distribución de probabilidad continua, por lo que se procedió a determinarla como una distribución discreta empírica como sigue.

Distribución Empírica Severidad				
Clases (x)		Frecuencia	P(x)	P(x<X)
\$	434.358	2	0,01342282	0,01342282
\$	542.385	118	0,79194631	0,80536913
\$	587.052	1	0,00671141	0,81208054
\$	593.433	2	0,01342282	0,82550336
\$	638.100	16	0,10738255	0,93288591
\$	1.084.770	9	0,06040268	0,99328859
\$	1.276.200	1	0,00671141	1
Total		149	1	

Tabla 8: Distribución empírica discreta de la severidad. Obtenido de: autoría propia

A continuación, para el procedimiento en cuestión 9073 se procedió a determinar la cantidad de simulaciones a realizar mediante la relación estrecha entre el percentil indicado y la función de pérdidas. Para este percentil, la cantidad N de simulaciones a realizar es de N=1000.

Se procedió a simular 1000 veces la función de pérdidas con las entradas de una distribución de Poisson con media 12,42 mediante la función RiskPoisson(12,42) como representación de la frecuencia y la severidad mediante una RiskDiscrete teniendo como valores x las distintas posibilidades del costo del procedimiento y como probabilidades registradas en la distribución empírica. La función de pérdidas LPA se calculó mediante la suma de los costos y añadiendo la función RiskOutput. Finalmente se definió el OpVar como el percentil %99,9 del LPA. El resultado según la simulación indicó que el OpVar es de \$15.703.641.

El comportamiento de los datos es como se muestra a continuación.

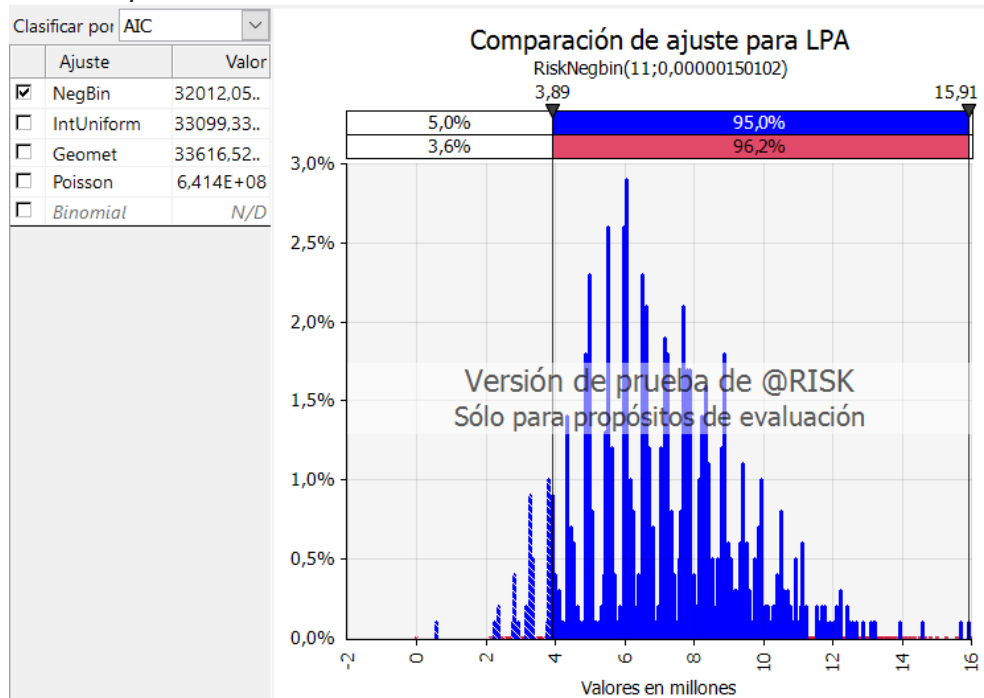


Figura 33: Ajuste de LPA según Risk Utilizando AIC. Obtenido de: autoría propia

Para el procedimiento mencionado del 12211 se replicó la metodología anterior. Los histogramas, estadística descriptiva y ajuste fueron como sigue:

descriptive statistics	
data points	12
minimum	0.
maximum	26.
mean	6.33333
median	4.5
mode	13.
standard deviation	8.32666
variance	69.3333
coefficient of variation	131.474
skewness	1.58216
kurtosis	1.10015

Figura 34: Estadística descriptiva frecuencia procedimiento 12211 analizado. Obtenido de: autoría propia

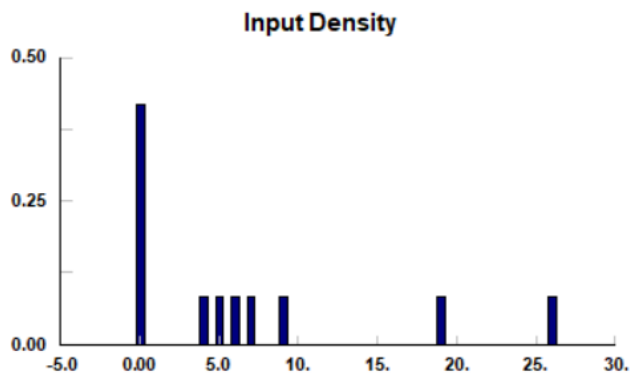


Figura 35: Histograma Frecuencia 12211. Obtenido de: autoría propia

Como se observa con base en el histograma, la frecuencia tiene un decaimiento exponencial marcado, aspecto típico de una distribución geométrica. A pesar de ser una distribución que mide la cantidad de fallas antes de un éxito, sirve para modelar los eventos en un período de tiempo n .

Auto::Fit of Distributions		
distribution	rank	acceptance
Geometric(0.136)	100	do not reject
Poisson(6.33)	8.75	reject
Discrete Uniform(0., 26.)	2.83	reject

Figura 36: Medida de ajuste de distribución geométrica. Obtenido de: autoría propia

De acuerdo con lo observado la distribución ajustada resultó ser una Geométrica. A pesar de no ser el comportamiento ideal de Poisson buscado este comportamiento sirve para representar las tasas de llegada como frecuencia.

Por otro lado, al observar el comportamiento de la severidad se observa que todos los datos tienen el mismo costo excepto uno, es decir 75 datos tienen un costo de \$411.700 y uno tiene un costo de \$823.400. Con base en esto, ajustar la distribución de severidad para el 12211 es indicado ajustar a una distribución discreta empírica.

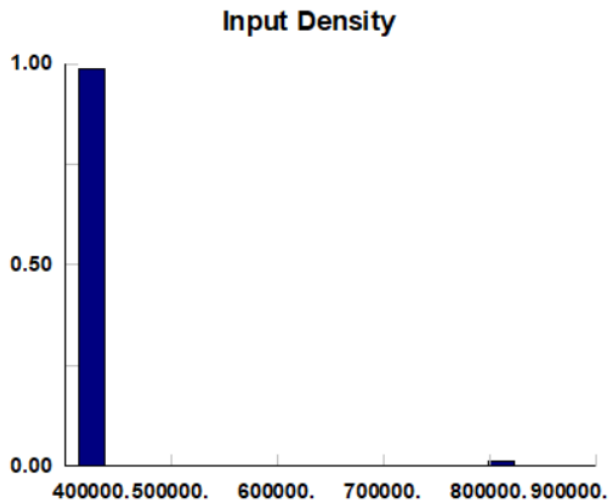


Figura 37: Histograma procedimiento 12211. Obtenido de: autoría propia

Auto::Fit of Distributions

distribution	rank	acceptance
Beta[4.12e+005, 8.34e+005, 1.03e-003, 7.44e-0	0.	reject
Erlang[4.12e+005, 1., 5.42e+003]	0.	reject
Exponential[4.12e+005, 5.42e+003]	0.	reject
Gamma[4.12e+005, 1., 5.42e+003]	0.	reject
Inverse Gaussian	no fit	reject
LogLogistic[4.12e+005, 0.944, 1.]	0.	reject
Lognormal[4.12e+005, 12.9, 3.49e-016]	0.	reject
Pareto[4.12e+005, 110]	0.	reject
Pearson 5[4.12e+005, 1.43e+015, 5.91e+020]	0.	reject
Pearson 6[4.12e+005, 218, 122, 6.36e-002]	0.	reject
Triangular[4.12e+005, 8.29e+005, 4.12e+005]	0.	reject
Uniform[4.12e+005, 8.23e+005]	0.	reject
Weibull[4.12e+005, 1., 4.12e+005]	0.	reject
Johnson SB	no fit	reject
Rayleigh[4.12e+005, 2.91e+005]	0.	reject
Inverse Weibull[4.12e+005, 85.9, 0.]	0.	reject
Chi Squared	no fit	reject
Power Function	no fit	reject

Figura 38: Prueba de bondad de Ajuste para severidad. Obtenido de: autoría propia

Como se había supuesto, la distribución de severidad para el procedimiento 12211 no se ajusta a ninguna distribución continua por lo que establecer una distribución empírica discreta es un camino prudente a optar.

Para calcular a definir la frecuencia para el modelo se definió una función RiskGeomet(0,136) y la severidad se modeló una distribución empírica discreta como se muestra a continuación.

<u>Otras entradas Conocidas</u>			
Distribución Empírica Costo			
<i>Clases</i>	<i>Frecuencia</i>	<i>P(x)</i>	<i>P(x<X)</i>
411700	75	0,98684211	0,98684211
823400	1	0,01315789	1
	76		

Figura 39: Distribución Discreta Empírica. Obtenido de: autoría propia

Con los datos de frecuencia como los valores x de una función RiskDiscrete y los valores de probabilidad aquellos denominados P(x).

De nuevo se eligió la función RiskOutput con la suma de los costos y se procedió a realizar 1000 simulaciones para el nivel sugerido del percentil del LPA.

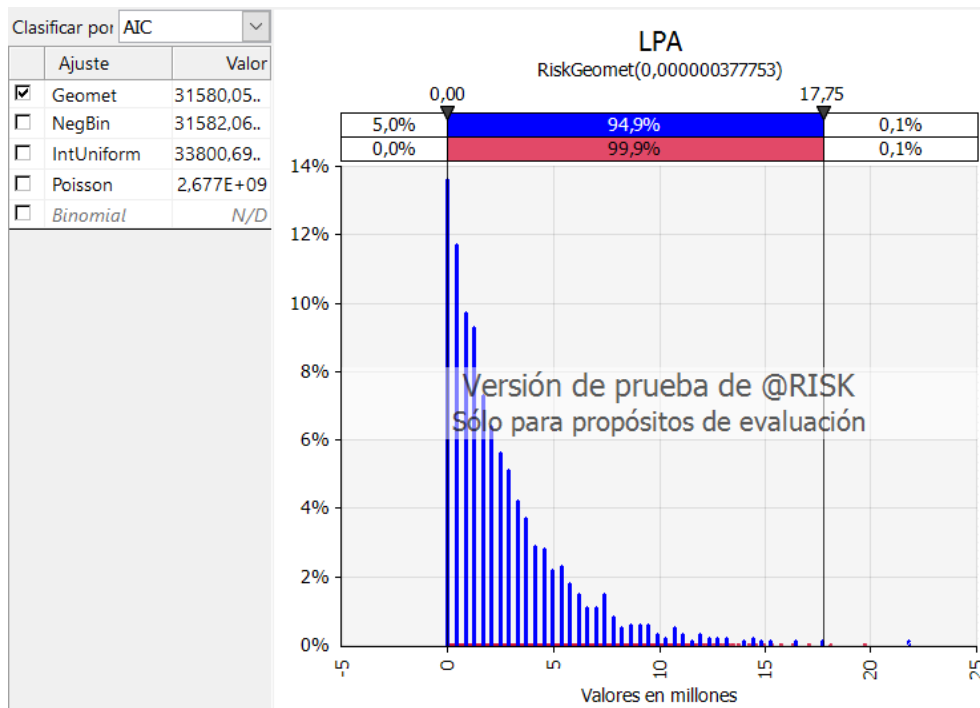


Figura 40: Ajuste LPA según Risk Utilizando AIC para 12211. Obtenido de: autoría propia

Mediante la simulación el monto sugerido del OpVar es de \$17.703.100.

3. PRESENTACIÓN Y DISCUSIÓN DE RESULTADOS

Estadística Descriptiva Muestra:

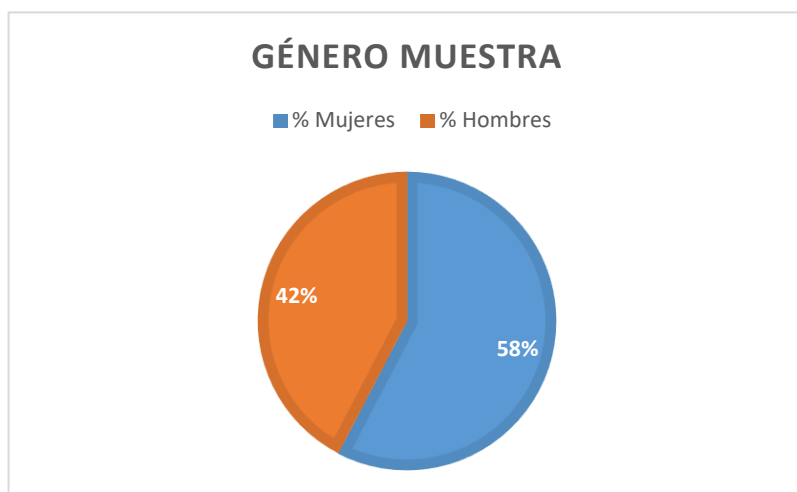


Figura 41: Muestra el porcentaje de cada género en la muestra poblacional. Fuente: Autoría propia basada en datos RIPS.

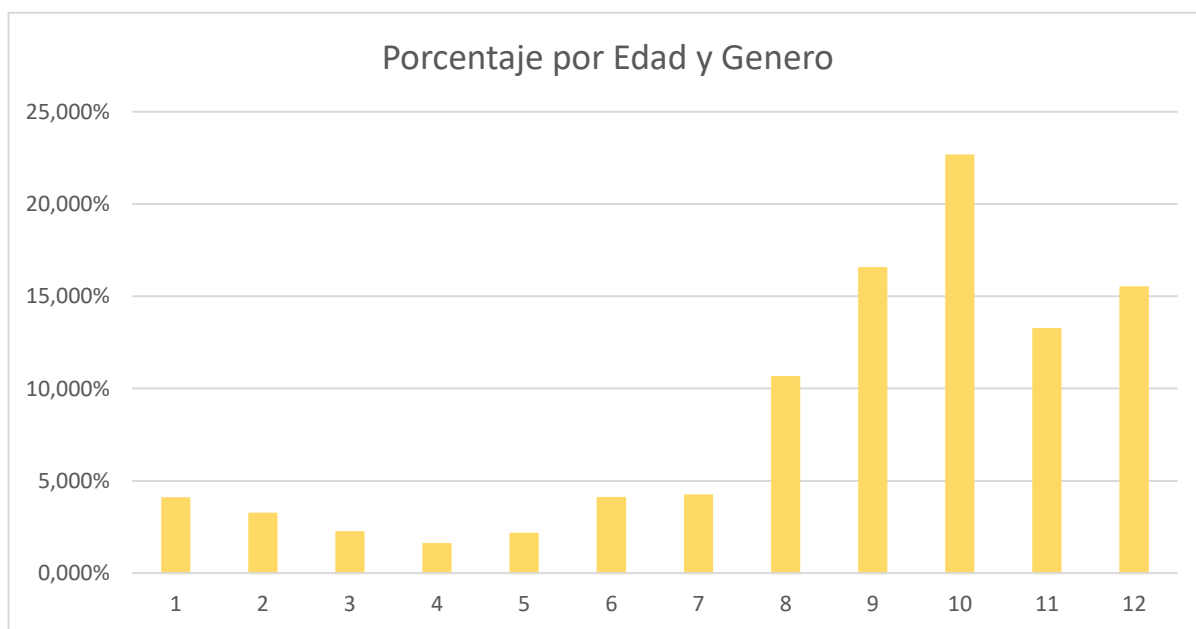


Figura 42: Muestra el porcentaje de personas atendidas por grupo etario y género (referirse a la tabla a). Fuente: Autoría propia basada en datos RIPS

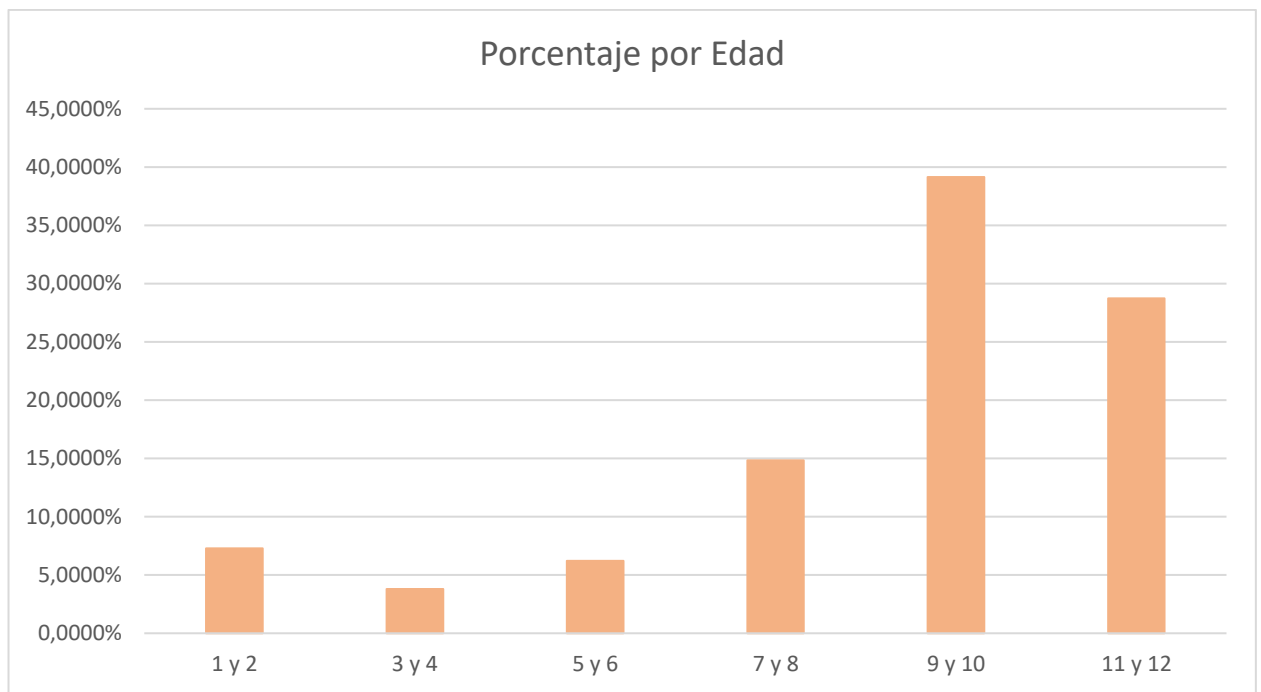


Figura 43: Muestra el porcentaje de gente atendida por ciclo vital (referencia a tabla a). Fuente: Autoría propia basada en datos RIPS.

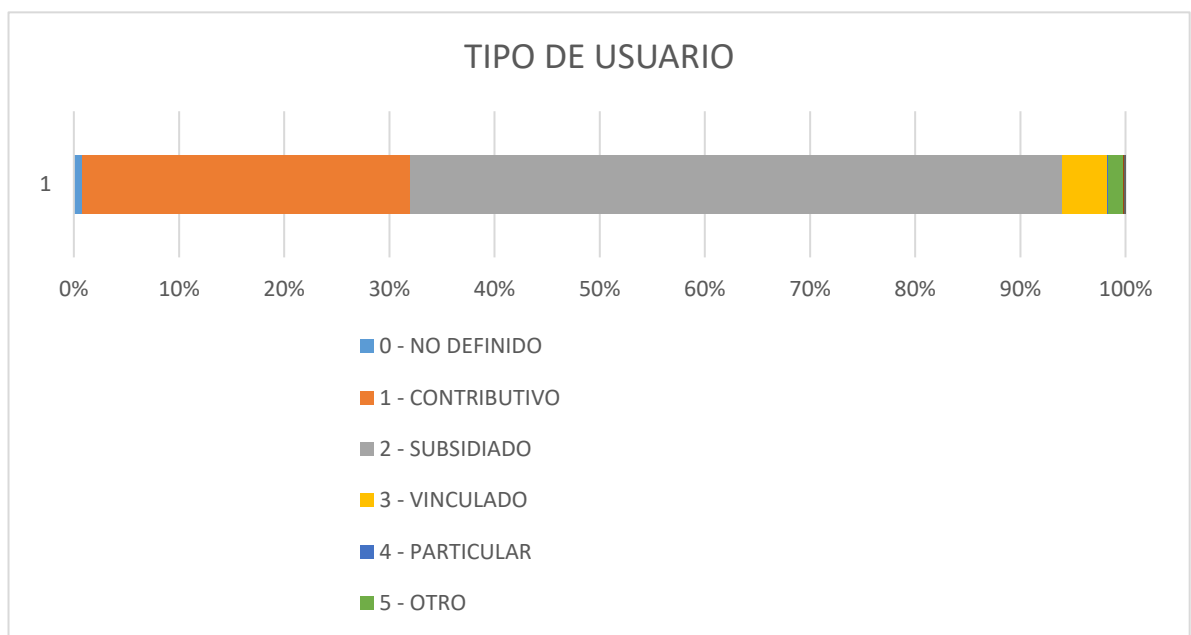


Figura 44: Muestra el porcentaje del total de cada Tipo de Usuario, las tres categorías más importantes son el Contributivo, el Subsidiado y el vinculado. Las otras categorías son casi nulas. Fuente: Autoría propia basada en datos RIPS.

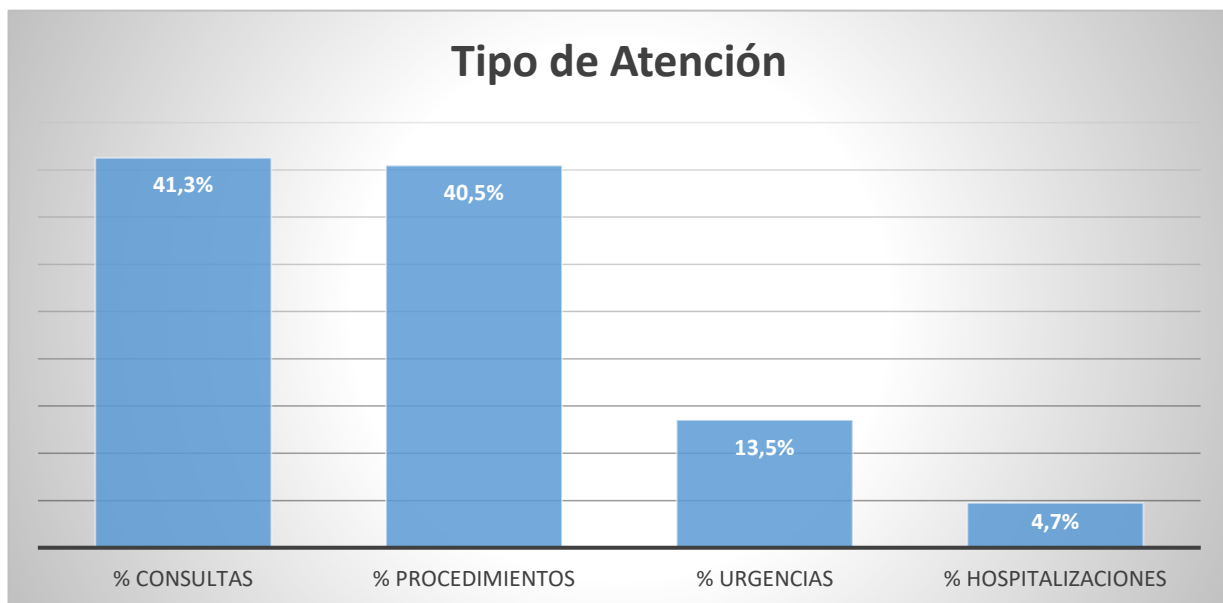


Figura 45: Muestra el Tipo de Atención que se realizó a los pacientes en estudio. Fuente: Autoría propia basada en datos RIPS.

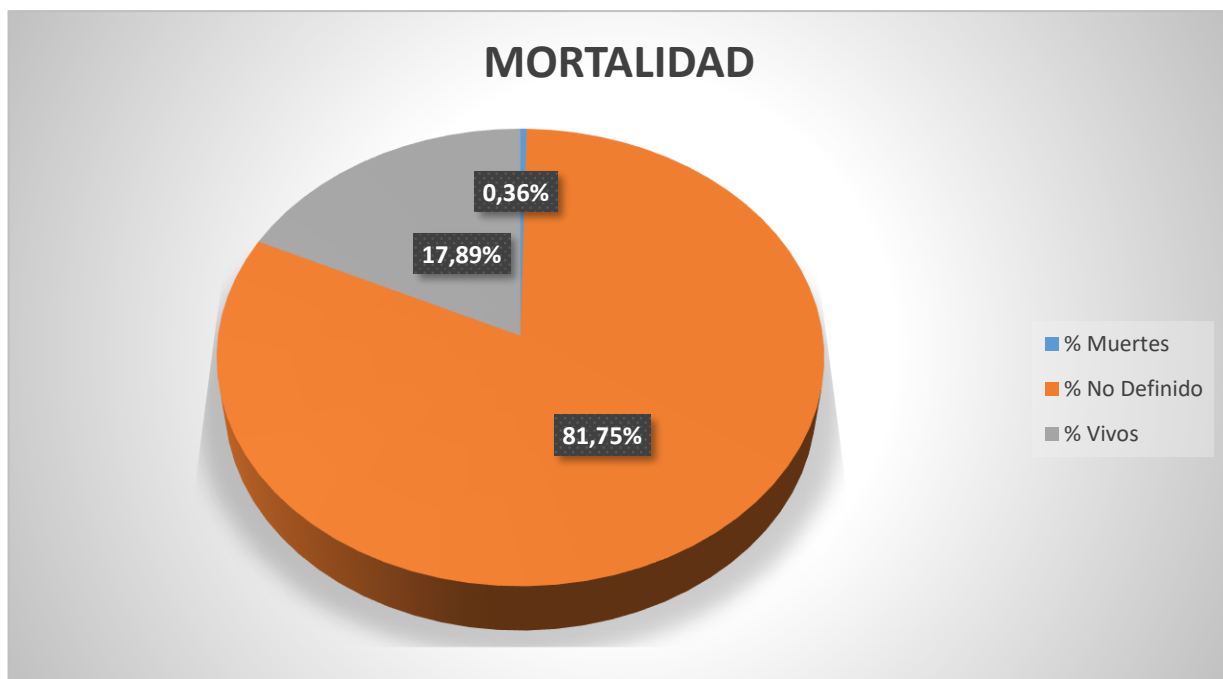


Figura 46: Mortalidad o Resultados de salida después de intervención o visita médica. Fuente: Autoría Propia basada en datos RIPS.

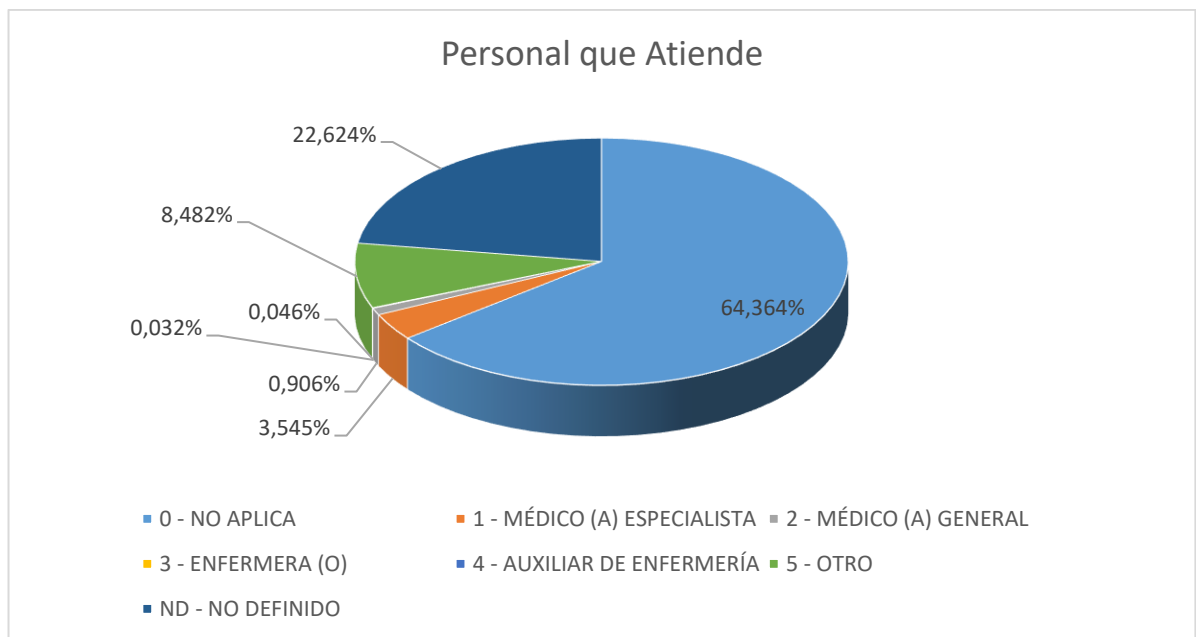


Figura 47: Muestra el porcentaje de personal que atiende a todas la personas que llegan al Hospital General de Medellín. Fuente: Autoría propia basado en los datos RIPS.

Días Estancia		Histograma Días Estancia		
		Clase	Frecuencia	% acumulado
Media	0.3976	0	826279	95.50%
Error típico	0.0047	2	13806	97.10%
Mediana	-	4	7115	97.92%
Moda	-	6	4710	98.46%
Desviación estándar	4.3466	8	3172	98.83%
Varianza de la muestra	18.8932	10	2164	99.08%
Curtosis	7,009	12	1434	99.24%
Coefficiente de asimetría	51.9172	14	1135	99.38%
Rango	1,086	16	826	99.47%
Mínimo	-	18	607	99.54%
Máximo	1,086	20	540	99.60%
Suma	344,007	22	391	99.65%
Cuenta	865,217	y mayor...	3038	100.00%

Tablas 9: Muestran la estadística descriptiva para los días de estancia en el Hospital General de Medellín. El histograma muestra una distribución semejante a una curva exponencial con una altísima concentración en el 0. Fuente: Datos RIPS

En las gráficas y tablas de estadística descriptiva presentadas anteriormente, se puede realizar una caracterización de la población que frecuenta los hospitales nivel III en Medellín. Adicionalmente, brinda algunas percepciones para comenzar a analizar la morbilidad en los datos de estudio. En la gráfica 1 se observa que la mayoría de la población que frecuenta estos servicios de salud es mujer. Observando un poco los datos se puede evidenciar que sucesos relacionados con los partos y diferencias hormonales llevan a las mujeres a visitar en una mayor medida los servicios de salud.

En las gráficas 2 y 3, se evidencia también que el porcentaje de los grupos 8 a 12 es mayor a todos los otros. Lo anterior, muestra que quienes frecuentan en su mayoría los servicios de salud son personas en su adultez o vejez. Mientras mayor es la edad, más propensas se vuelven las personas a problemas de salud.

La gráfica 5 pudiera ser de gran ayuda para los hospitales nivel III en Medellín. Esta, muestra la proporción del número de atenciones para cada tipo de atención en los hospitales. Se puede observar como la mayoría de las personas requieren servicios de consultas procedimientos, los cuales pueden ser llevados a cabo en su mayoría por personal menos capacitado (no especialistas) ya que en la mayoría los casos no son de urgencia y pueden remitirlos para segundas opiniones. Sería interesante evaluar el impacto que tiene cada una de las atenciones en los costos del hospital para poder enfocar esfuerzos y lograr costos más competitivos en los hospitales.

La gráfica 6 y 7 evidencian los grandes problemas de realizar estudios con base en información pública en Colombia. Se considera inaudito que los datos públicos de registros individuales de salud (dónde probablemente provendrán todas las investigaciones importantes y en base a estas realizar programas y enfocar presupuesto) sean así de incompletos. Observamos como casi el 82% de los datos de mortalidad hacen referencia a No Definido. No es posible que los hospitales no sepan si después de realizar cualquier tipo de intervención el paciente se encuentra vivo o muerto. Es increíble además que Colombia no tenga la información clara de que porcentaje de las atenciones las presta un médico general, un médico especialista, una enfermera o cualquier otro; pues un 64% de los campos en los RIPS están llenos con no aplica para la persona que se encargó principalmente de ejecutar el procedimiento. Se encontraron varios campos como este que no brindan ninguna certeza en los datos para realizar un análisis.

Por último, se estudió un dato que podría ser de vital importancia en el estudio de capacidad para los hospitales nivel III. Este dato hace referencia a los días de permanencia de una persona promedio en el hospital. La media de los días de estancia por persona tiene una media de 0.3976 y una desviación estándar de 4.35. Encontramos además un máximo de 1086 días de estancia, aunque el percentil 96 se encuentre lleno de 0 días de estancia, se debe profundizar el estudio de casos como este para ver cómo se pueden volver más eficientes y si hay una forma de

disminuir la cantidad de días de estancia o incrementar la capacidad instalada de los hospitales.

Resultados Morbilidad:

Enfermedad	Conteo	Porcentaje
Z139 - EXAMEN DE PESQUISA ESPECIAL, NO ESPECIFICADO	81955	9,47
R688 - OTROS SINTOMAS Y SIGNOS GENERALES ESPECIFICADOS	48966	5,66
Z000 - EXAMEN MEDICO GENERAL	45000	5,20
N189 - INSUFICIENCIA RENAL CRONICA, NO ESPECIFICADA	26579	3,07
R104 - OTROS DOLORES ABDOMINALES Y LOS NO ESPECIFICADOS	19587	2,26
N390 - INFECCION DE VIAS URINARIAS, SITIO NO ESPECIFICADO	17190	1,99
R509 - FIEBRE, NO ESPECIFICADA	9562	1,11
O800 - PARTO UNICO ESPONTANEO, PRESENTACION CEFALICA DE VERTICE	9427	1,09
R51X - CEFALEA	8954	1,03
OTROS	597997	69,12

Tabla 10: Tabla que muestra los resultados de morbilidad general para toda la muestra poblacional, esto hasta la frecuencia del 1% de las personas atendidas en total. Obtenido de: autoría propia basada en datos RIPS

En la tabla 10 se presenta una porción del resumen general de morbilidad de los datos. A partir de esta tabla se pueden observar las enfermedades relevantes hasta el 1% de frecuencia del total general. Como se observa, esta primera tabla de resultados permite visibilizar muy poco la verdadera morbilidad de la población, esto debido a que muchas enfermedades que son descritas por los datos como “general o no especificado” no revelan más información sobre la enfermedad padecida por el paciente. Investigando más a profundidad, con el diagnóstico Z139 se identificó además que el 100% de los diagnósticos de egreso eran “No reportado” dejando aún más incógnitas sobre la morbilidad de una gran fracción de la población. Se define, el diagnóstico Z139 por ejemplo, Examen de Pesquisa como: “Aquellos casos en que ciertas circunstancias que no son enfermedades, lesiones ni causas externas clasificables en las categorías A00–Y89, se registran como “diagnósticos” o “problemas”. Esto puede surgir principalmente de dos maneras:

(a) -Cuando una persona que puede o no estar enferma en ese momento, entra en contacto con los servicios de salud para algún propósito específico, tal como recibir una atención o servicios limitados por una condición presente, donar un órgano o tejido, recibir una vacuna o discutir un problema a que no es por sí mismo una enfermedad o lesión.

(b) -Cuando existe alguna circunstancia o problema que influye en el estado de salud de una persona, pero no es en sí misma una enfermedad o lesión actual. Estos factores pueden descubrirse en encuestas en la población, cuando la persona

puede o no estar enferma, o registrarse como un factor adicional a ser tomado en cuenta cuando la persona reciba atención por alguna enfermedad o lesión” (WAO, 2016)

Según esta definición se puede deducir que no necesariamente todos aquellos pacientes que entran a hospitales nivel III padecen una enfermedad y es posible que el diagnóstico más recurrente haga referencia en ocasiones a estos casos, o por otro lado que la información es incompleta y perseguir el resultado de estos diagnósticos y conocer su resultado permitiría identificar realmente estos que enfermedades presentan estos individuos.

Aunque una parte de la tabla hace referencia a enfermedades no específicas, brinda información importante acerca de las enfermedades N189, N390, O800 y R51X las cuales pudieran ser clasificadas como las enfermedades más recurrentes y por ende más importantes a analizar posteriormente para los hospitales nivel III. Sería de gran importancia conocer el resultado de mortalidad para estas enfermedades y poderlas clasificar por su criticidad. Se analiza, posteriormente, la clasificación de morbilidad por grupos poblacionales preestablecidos.

Grupo	Enfermedad	Porcentaje
1	Z139 - EXAMEN DE PESQUISA ESPECIAL, NO ESPECIFICADO	9.31
1	Z000 - EXAMEN MEDICO GENERAL	8.18
2	Z139 - EXAMEN DE PESQUISA ESPECIAL, NO ESPECIFICADO	9.36
2	R688 - OTROS SINTOMAS Y SIGNOS GENERALES ESPECIFICADOS	4.98
3	Z139 - EXAMEN DE PESQUISA ESPECIAL, NO ESPECIFICADO	11.43
3	R688 - OTROS SINTOMAS Y SIGNOS GENERALES ESPECIFICADOS	5.87
4	Z139 - EXAMEN DE PESQUISA ESPECIAL, NO ESPECIFICADO	11.72
4	R688 - OTROS SINTOMAS Y SIGNOS GENERALES ESPECIFICADOS	7.62
5	Z139 - EXAMEN DE PESQUISA ESPECIAL, NO ESPECIFICADO	8.59
5	R688 - OTROS SINTOMAS Y SIGNOS GENERALES ESPECIFICADOS	4.69
6	R688 - OTROS SINTOMAS Y SIGNOS GENERALES ESPECIFICADOS	6.31
6	Z139 - EXAMEN DE PESQUISA ESPECIAL, NO ESPECIFICADO	5.65
7	Z139 - EXAMEN DE PESQUISA ESPECIAL, NO ESPECIFICADO	9.25
7	R104 - OTROS DOLORES ABDOMINALES Y LOS NO ESPECIFICADOS	2.62
8	Z139 - EXAMEN DE PESQUISA ESPECIAL, NO ESPECIFICADO	6.98
8	O800 - PARTO UNICO ESPONTANEO, PRESENTACION CEFALICA DE VER	5.53
9	Z139 - EXAMEN DE PESQUISA ESPECIAL, NO ESPECIFICADO	8.82
9	Z000 - EXAMEN MEDICO GENERAL	4.74
10	Z139 - EXAMEN DE PESQUISA ESPECIAL, NO ESPECIFICADO	10.65
10	R688 - OTROS SINTOMAS Y SIGNOS GENERALES ESPECIFICADOS	6.74
11	Z139 - EXAMEN DE PESQUISA ESPECIAL, NO ESPECIFICADO	11.97
11	Z000 - EXAMEN MEDICO GENERAL	7.37
12	N189 - INSUFICIENCIA RENAL CRONICA, NO ESPECIFICADA	8.95
12	Z139 - EXAMEN DE PESQUISA ESPECIAL, NO ESPECIFICADO	8.78

Tabla 11: Muestra los resultados de la segmentación por grupos (tabla a) con las dos enfermedades más recurrentes en cada grupo y sus porcentajes respectivos. Fuente: Autoría propia. Obtenido de: autoría propia basada en datos RIPS

La tabla anterior (11) permite realizar una caracterización más a profundidad de la morbilidad. Esta, arroja el resultado de morbilidad referente a grupos poblacionales específicos. Se puede observar a partir de esta tabla, como las enfermedades observadas “generales o no específicas” son transversales para todos los grupos poblacionales. Por ejemplo, la enfermedad Z139 que fue clasificada como la más recurrente para todos los datos, es transversal a todos los grupos poblacionales pues se ve clasificada en alguno de los primeros dos puestos para todos los grupos. Por otra parte, se observa que las enfermedades clasificadas como específicas en la clasificación general se conectan principalmente con uno de los grupos poblacionales establecidos. Esto se puede observar con la enfermedad O800 que se conecta con la segunda enfermedad más frecuente del grupo 8 y la N189 que se conecta con la enfermedad más recurrente en el grupo 12. Esta información pudiera ser de gran relevancia para los servicios de salud, pues demuestra una conexidad de las enfermedades recurrentes con un grupo poblacional específico. Con base en esto, se desarrollan análisis específicos de la enfermedad Z139 y la O800 (una general y una específica, respectivamente).

Enfermedad Z139:

Municipio Residencia	ConteoMunicipio	Descripción
12086	29432	MEDELLÍN
12659	11590	QUIBDÓ
11706	6075	EL CARMEN DE ATRATO
12292	3653	ISTMINA
11803	2073	TADÓ
12148	2011	RIOSUCIO
12385	1779	ALTO BAUDÓ
11942	1722	BOJAYA
11992	1684	ENVIGADO
11729	1682	BELLO

Tabla 12: Frecuencia enfermedad Z139 por municipio de procedencia del paciente. Obtenido de: autoría propia basada en datos RIPS

Se observa de la clasificación por municipio resultados muy concretos de donde se concentran los mayores focos de enfermedades (proveniencia de las personas). Se observa que la mayor cantidad de pacientes con este diagnóstico provienen de Medellín seguidas del Chocó. Para hospitales nivel III de Medellín se espera que gran parte de la población atendida sea del mismo municipio, pero se puede observar que hay gran procedencia de zonas que no cuentan con el mismo nivel de atención de salud.

La clasificación ABC de los procedimientos referentes a este diagnóstico se puede visualizar a continuación:

ID Proc	Conteo	CostoProcedimiento	PorcentajeID	PorcentajeCostoAcumulado	Clasificación	Descripción
8692	42	\$ 346,925,935	0.15	5.42	A	599500 - LITIASIS URINARIA FRAGMENTADA INTRACORPOREA ENDOSCÓPICA SOD +
12362	133	\$ 305,951,940	0.30	10.21	A	883101 - RESONANCIA NUCLEAR MAGNETICA DE CEREBRO
12211	557	\$ 286,778,710	0.45	14.69	A	879111 - TOMOGRAFIA AXIAL COMPUTADA DE CRANEO SIMPLE
12371	121	\$ 269,168,410	0.60	18.90	A	883230 - RESONANCIA NUCLEAR MAGNETICA DE COLUMNA LUMBOSACRA SIMPLE +
12247	442	\$ 264,494,800	0.75	23.03	A	881234 - ECOCARDIOGRAMA MODO M Y BIDIMENSIONAL CON DOPPLER A COLOR
12229	347	\$ 232,293,950	0.90	26.67	A	879420 - TOMOGRAFIA AXIAL COMPUTADA DE ABDOMEN Y PELVIS (ABDOMEN TOTAL)
13587	3836	\$ 198,020,336	1.05	29.76	A	906913 - PROTEÍNA C REACTIVA
13712	277	\$ 197,888,620	1.20	32.85	A	911107 - PROCESAMIENTO DE LA UNIDAD DE GLÓBULOS ROJOS O ERITROCITOS LEUCORE
12838	6708	\$ 167,971,445	1.35	35.48	A	902210 - HEMOGRAMA IV [HEMOGLOBINA
12226	297	\$ 146,623,250	1.50	37.77	A	879301 - TOMOGRAFIA AXIAL COMPUTADA DE TORAX
13036	3378	\$ 127,033,180	1.65	39.76	A	903859 - POTASIO +
12621	101	\$ 106,812,220	1.80	41.43	A	898103 - ESTUDIO DE COLORACIÓN INMUNOHISTOQUIMICA EN BIOPSIA +
12383	45	\$ 103,683,540	1.95	43.05	A	883401 - RESONANCIA NUCLEAR MAGNETICA DE ABDOMEN +
13018	4827	\$ 100,292,648	2.10	44.62	A	903841 - GLUCOSA EN SUERO
12007	1335	\$ 98,242,830	2.25	46.15	A	871121 - RADIOGRAFÍA DE TORAX (P.A. O A.P Y LATERAL
13041	3167	\$ 98,142,000	2.40	47.69	A	903864 - SODIO+
12954	296	\$ 81,687,922	2.55	48.97	A	903437 - TROPONINA I
12739	79	\$ 72,188,382	2.70	50.09	A	901314 - Mycobacterium

Tabla 13: Fragmento de la clasificación ABC de los procedimientos de la enfermedad Z139 para la población general. Obtenido de: autoría propia basada en datos RIPS

En la tabla anterior, se observan los procedimientos clasificados (para la enfermedad Z139) por el costo total más elevado para los años de estudio. Estos procedimientos podrían dar más información acerca de la verdadera sintomatología de los pacientes en cuestión si se conociera su resultado, aspecto clave a tener en consideración.

La clasificación ABC permite al hospital conocer sus procedimientos más costosos por enfermedad. En este caso, se observa que el procedimiento 8692, referente a litiasis urinaria, es el más costoso con un valor de \$346.925.935. Es en este tipo de casos donde los hospitales o entidades de servicios de salud debieran analizar los costos totales de servicio para tratar de reducirlos, teniendo en cuenta el tipo de procedimiento a realizar. Este estudio debería ser llevado a cabo para todos los resultados clasificados como A en el ABC de procedimientos y replicado para cada enfermedad importante.

Enfermedad O800:

Municipio	Residencia	Conteo	Municipio	Descripción
	12086		3064	MEDELLÍN
	11729		346	BELLO
	12644		221	ITAGUI
	12222		80	NECHÍ
	12659		75	QUIBDÓ
	11820		45	COPACABANA
	12428		42	SEGOVIA
	12223		36	SABANETA
	11992		31	ENVIGADO
	11772		31	ZARAGOZA

Tabla 14: Frecuencia enfermedad O800 por municipio de procedencia del paciente. Obtenido de: autoría propia basada en datos RIPS

Para la enfermedad O800 se encuentra un patrón parecido al de la enfermedad Z139. Las personas provenientes de Medellín ocupan el principal foco de visitas debido a esta condición y esta vez se observa como municipios cercanos, como Bello e Itagüí, utilizan los servicios de hospitales nivel III en Medellín. Se observa un foco más importante que el de Chocó en el Bajo Cauca antioqueño.

La clasificación ABC de los procedimientos referentes a este diagnóstico se puede visualizar a continuación:

ID	Proced	Conteo	Costo	Procedimiento	Porcentaje	Costo	Porcentaje	ID	Procedimiento	Costo	Acumulado	Clasificación	Descripción
9073	890	\$	510.783.089		61,89	0,98		61,89	A				735300 - ASISTENCIA DEL PARTO NORMAL CON EPISIOTOMIA Y/O PERINEORRAFIA SOD
9074	434	\$	249.512.976		30,23	1,96		92,13	A				735910 - ASISTENCIA DEL PARTO ESPONTANEO NORMAL (EXPULSIVO)
1	27	\$	13.685.085		1,66	2,94		93,79	A				1 - NO DEFINIDO
9069	16	\$	11.054.791		1,34	3,92		95,13	A				721001 - PARTO INSTRUMENTADO CON FORCEPS O ESPATULAS DE VELASCO BAJOS
9078	4	\$	3.461.795		0,42	4,90		95,54	A				740100 - CESAREA SEGMENTARIA TRANSPERITONEAL SOD +
8864	7	\$	2.859.284		0,35	5,88		95,89	A				663100 - SECCIÓN Y/O LIGADURA DE TROMPAS DE FALLOPIO [CIRUGIA DE POMEROY] POR MINILAPAROTOMIA SOD +
12228	1	\$	2.119.000		0,26	6,86		96,15	A				879410 - TOMOGRAFIA AXIAL COMPUTADA DE ABDOMEN SUPERIOR
7994	1	\$	1.805.250		0,22	7,84		96,37	A				451301 - ESOFAGOGASTRODUODENOSCOPIA (EGD) DIAGNOSTICA O EXPLORATORIA SIN BIOPSIA
9102	8	\$	1.804.191		0,22	8,82		96,59	A				759101 - DRENAJE DE COLECCIÓN OBSTETRICA (DE EPISIOTOMIA O EPISIOTOMIA) EN PERINE POR INCISIÓN
12254	1	\$	1.669.000		0,20	9,80		96,79	A				881302 - ULTRASONOGRAFIA DE ABDOMEN TOTAL: HIGADO
8866	2	\$	1.634.850		0,20	10,78		96,99	A				664001 - SALPINGECTOMIA UNILATERAL TOTAL POR LAPAROTOMIA
13353	15	\$	1.151.902		0,14	11,76		97,12	A				906249 - VIH 1 Y 2
14293	6	\$	1.037.200		0,13	12,75		97,25	A				A20002 - ADQUISICION Y ADMINISTRACION DE INSUMOS Y BIOLOGICOS INCLUYE: DISTRIBUCION
12601	6	\$	990.847		0,12	13,73		97,37	A				897012 - MONITORIA FETAL INTRAPARTO

Tabla 15: Fragmento de la clasificación ABC de los procedimientos de la enfermedad O800 para el grupo 8 femenino. Obtenido de: autoría propia basada en datos RIPS

La tabla del ABC de procedimientos para el diagnóstico O800 muestra que el procedimiento más costoso fue el 9073, representando cerca 61% del costo para todos los procedimientos con costo superior a 0 aplicados para los años 2014-2017, con un valor de \$510.783.089. Este seguido de los procedimientos 9074 y el

procedimiento uno, representando cada uno costos de 30% y 1,65%, con valores de \$249.512.976 y \$13.685.085 respectivamente. Como tal los primeros tres procedimientos más costosos hacen referencia a los códigos CUPS, 735300- Asistencia del Parto Normal con Episiorrrafia y/o perineorrafia, 735910- asistencia del parto espontáneo normal (Expulsivo) y finalmente 1-No Definido. Con estas descripciones es posible identificar de nuevo un problema recurrente y es que el procedimiento que representa el tercer mayor costo para este diagnóstico es “No Definido”, lo cual no permite una identificación de las causas de los costos de los procedimientos demandados. Por otro lado, se logra evidenciar que los partos en general representan la gran mayoría de los costos para el grupo 8 con género femenino, siendo entonces esto un punto válido de análisis para el hospital respecto a la demanda de personal capacitado y de recursos necesarios para poder satisfacer esta demanda. Además de esto identificar los directores principales de costos de estos procedimientos y buscar formas, si posible, de volverlo más eficiente.

Resultados Series de Tiempo:

Enfermedades Totales:

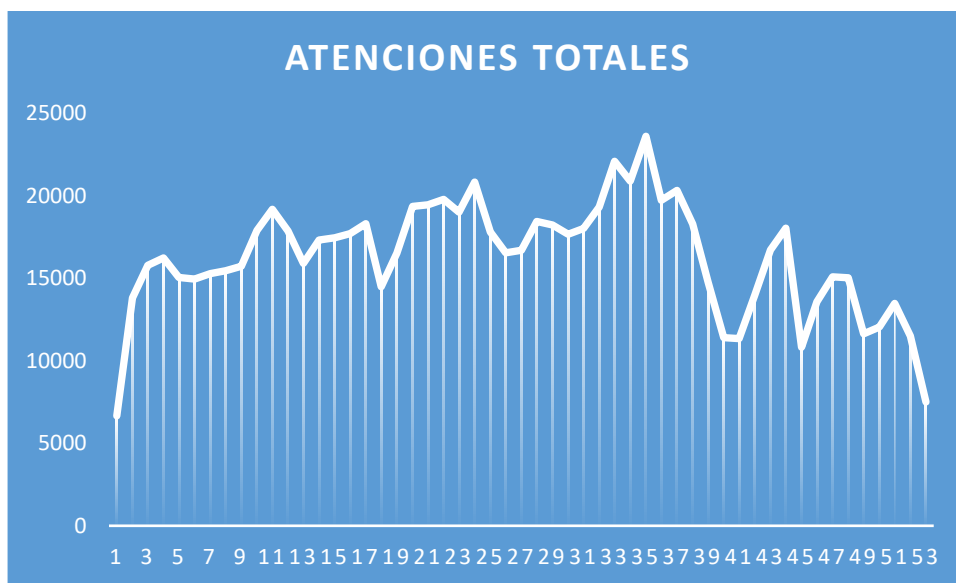


Figura 48: Gráfica de la serie de tiempo de los procedimientos totales durante los años del estudio por frecuencia semanal. Obtenido de: autoría propia basada en datos RIPS.

Con base en la serie de tiempo de procedimientos totales se evidencia que el número de atenciones totales durante las últimas y primeras semanas de año es históricamente bajo en comparación a las semanas en la mitad del

año. A continuación, se analizan diagnósticos individuales con sus respectivos procedimientos.

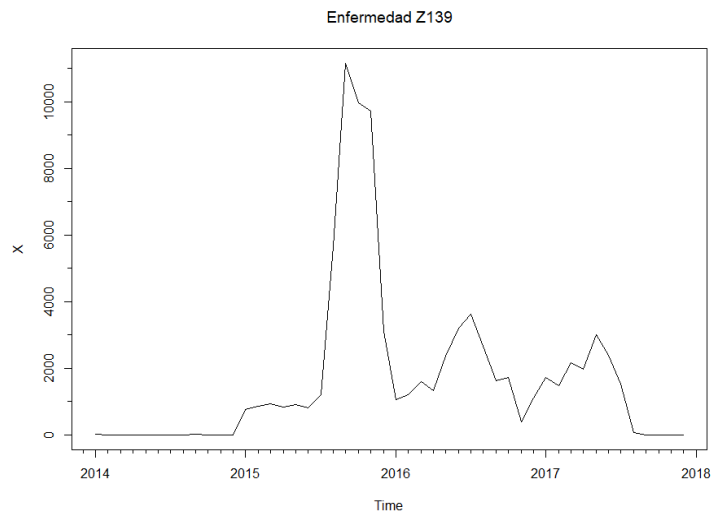


Figura 49: Serie de tiempo de la enfermedad Z139. Obtenido de: autoría propia basada en datos RIPS.

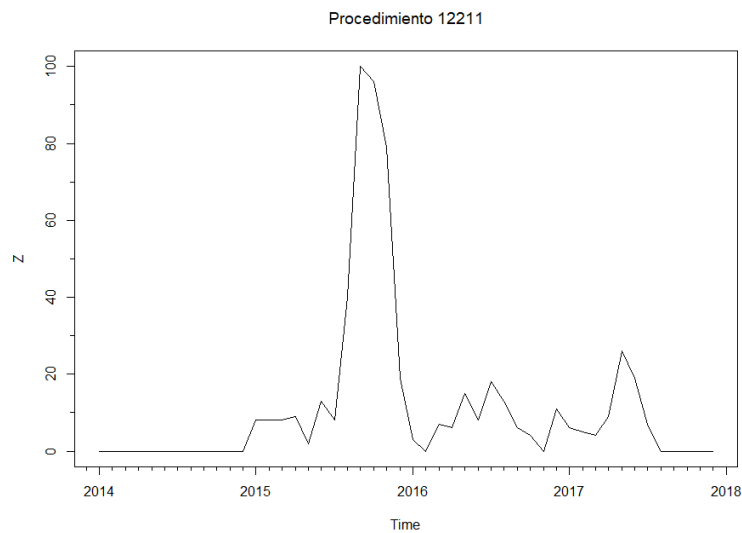


Figura 50: Serie de tiempo del procedimiento 12211, referente a la enfermedad Z139. Obtenido de: autoría propia basada en datos RIPS.

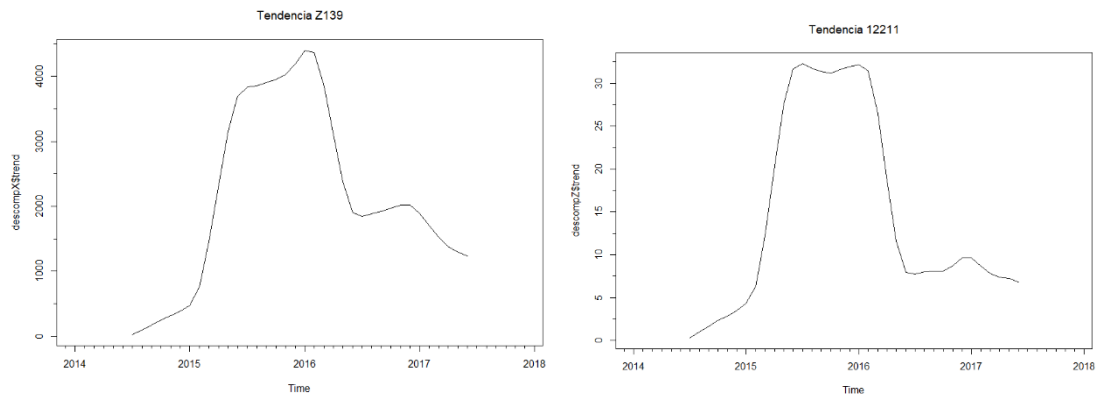


Figura 51: Descomposición de la componente de tendencia de la enfermedad Z139 y procedimiento 12211. Obtenido de: autoría propia basada en datos RIPS.

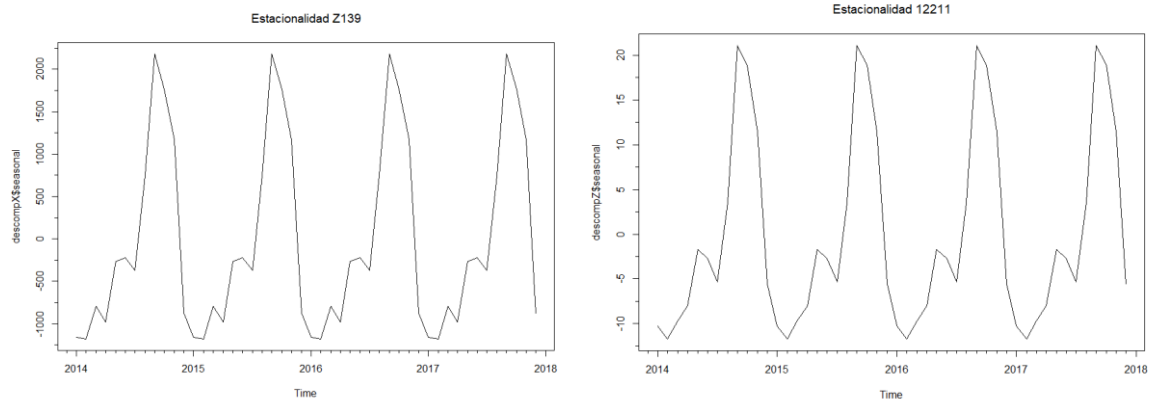


Figura 52: Descomposición de la componente estacional de la enfermedad Z139 y procedimiento 12211. Obtenido de: autoría propia basada en datos RIPS.

Enfermedad Z139 con Procedimiento 12211 (879111 - TOMOGRAFIA AXIAL COMPUTADA DE CRANEO SIMPLE): Se obtienen las siguientes gráficas en series de tiempo para dicha enfermedad y el respectivo procedimiento. Se observa claramente como la enfermedad y procedimiento están estrechamente relacionados, más aún el procedimiento aparenta tener un comportamiento muy similar al del diagnóstico, indicando la posible existencia de una correlación alta. Con respecto a la tendencia, se observa que esta varía a través del tiempo pero que presenta cambios bruscos cerca de los años 2015 y 2016. Este fenómeno puede ser de interés estudiarlo ya que presenta una variación muy grande y se puede deber a una causa asignable. Este evento se dio específicamente entre los meses de Julio y Agosto para el 2015, en donde el diagnóstico Z139 creció 10000 unidades, después de Agosto cayó en picada hasta el principio del 2016 a un nivel cercano al que tenía antes de este crecimiento abrupto.

En la componente estacional se observan picos bastante marcados en la mitad del año, especialmente durante el mes de Agosto. Esta información puede ser de gran interés para este tipo de hospitales ya que saben que deben prepararse para afrontar este diagnóstico Z139 y por consecuencia realizar procedimientos 12211. De esta forma, los hospitales nivel III sabrán que la capacidad de este procedimiento, número de especialistas para la enfermedad, número de equipos para el procedimiento e insumos disponibles para el procedimiento deben aumentar para los meses de mitad de año. Por otro lado se observa que cerca de los meses del principio del año se presentan los niveles más bajos de demanda tanto del procedimiento como de ocurrencia del diagnóstico.

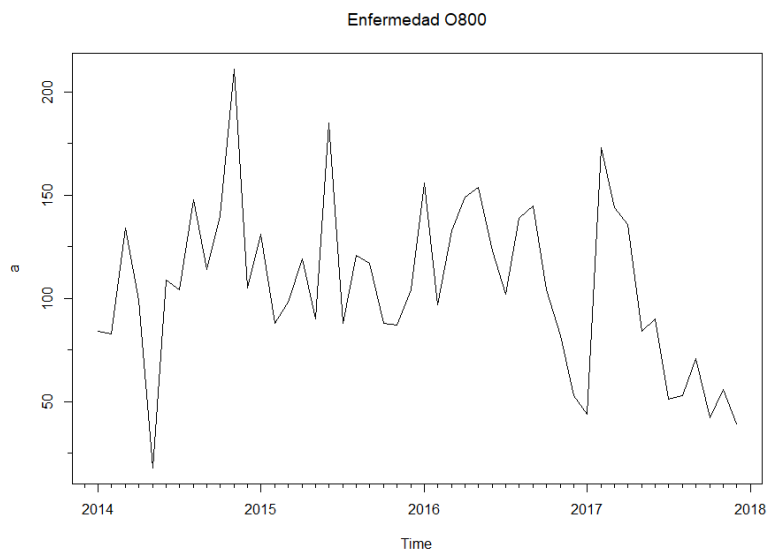


Figura 53: Serie de tiempo de la enfermedad O800. Obtenido de: autoría propia basada en datos RIPS.

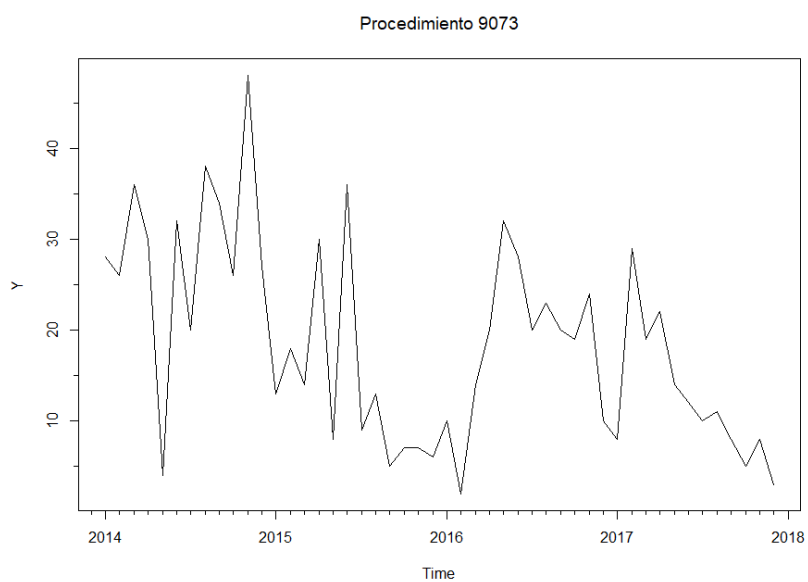


Figura 54: Serie de tiempo del procedimiento 9073, referente a la enfermedad O800. Obtenido de: autoría propia basada en datos RIPS.

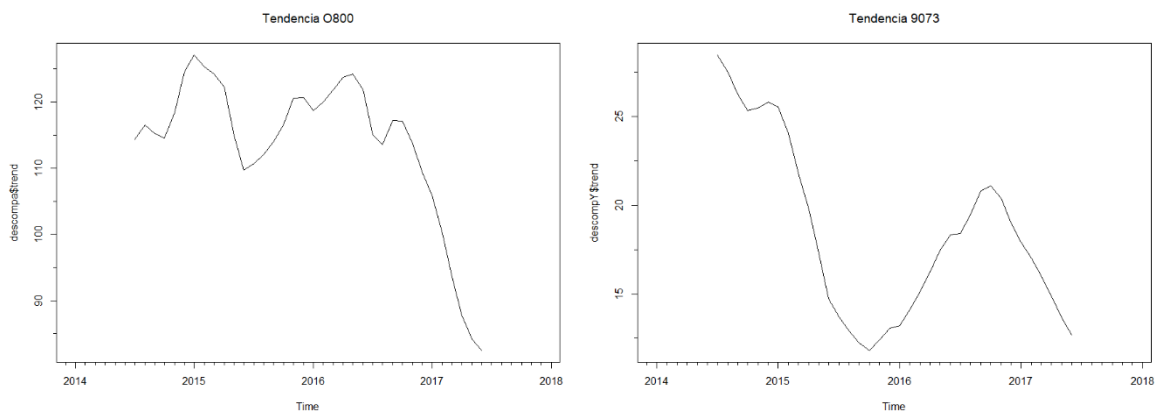


Figura 55: Descomposición de la componente de tendendencia de la enfermedad O800 y procedimiento 9073. Obtenido de: autoría propia basada en datos RIPS.

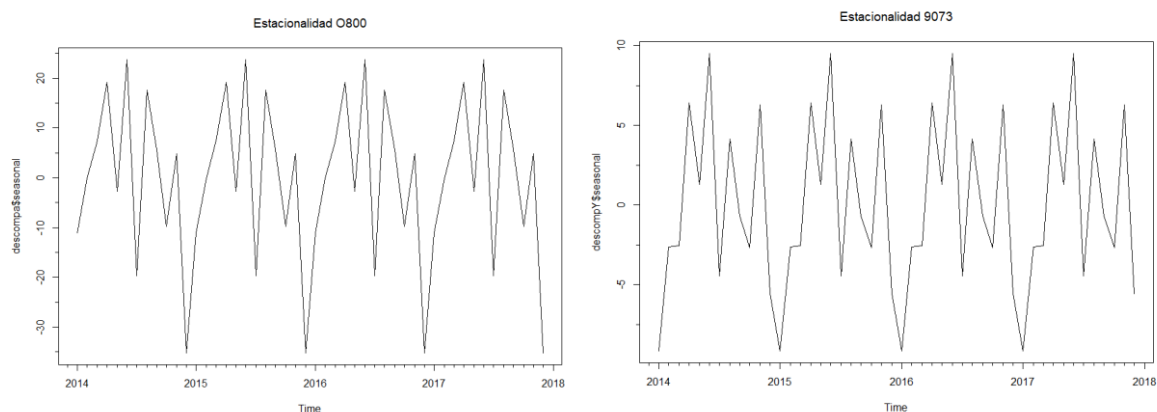


Figura 56: Descomposición de la componente estacional de la enfermedad O800 y procedimiento 9073. Obtenido de: autoría propia basada en datos RIPS.

Enfermedad O800 y Procedimiento 9073(735300- ASISTENCIA DEL PARTO NORMAL CON EPISIORRAFIA Y/O PERINEORAFIA): Se observa a partir de las dos gráficas del comportamiento de la enfermedad y el diagnóstico tanto similitudes como diferencias. Existen similitudes en los patrones de los períodos de Abril del 2014 con una gran declive, unos picos marcados de incrementos cerca de Noviembre del 2014, por otro lado se percibe de nuevo un declive cerca a marzo del 2017, pero tienen pocas coincidencias en muchas de las fechas restantes. En la descomposición de la tendencia se encuentra un comportamiento divergente entre los años 2015 y 2016; mientras la enfermedad O800 obtiene tendencias fuertes al alza, el procedimiento 9073 encuentra niveles de tendencia mucho más bajos y viceversa. Esto puede significar una correlación negativa entre ambas variables o una “correlación” con una dispersión de tiempo, es decir, las curvas pueden estar correlacionadas pero en períodos diferentes del tiempo (esto por tratarse de enfermedades relacionadas con partos).

Por otra parte, en la estacionalidad, se observan similitudes más estrechas demostrando picos bajos de demanda cerca de los meses de diciembre y enero y para ambos el pico mayor cerca de los meses Abril y Marzo.

Análisis procedimientos VAR:

ID Procedimiento 9073 (735300- ASISTENCIA DEL PARTO NORMAL CON EPISIORRAFIA Y/O PERINEORAFIA):

Para el procedimiento con ID 9073, referente a la asistencia de parto se evidencia según el resultado del OpVar que el capital mínimo requerido a tener para cada mes es de \$15.703.641. Este valor en teoría representa el capital que se requiere

para poder tener solvencia y pagar por ejemplo la mano de obra de especialistas o pagar los insumos requeridos para los procedimientos entre otros costos incurridos en caso de un evento extremo. Este valor percentil %99,9 de la función de pérdidas agregada. A pesar de esto, recomendar a un hospital nivel III en Colombia que tenga este capital con la actual crisis financiera que vive el sector de salud puede parecer una recomendación insensata, pero es un número que indica un capital sano y puede hasta agregar más información respecto a las necesidades de los institutos prestadores de servicios y los puntos que los hacen vulnerables.

Por otra parte, se observa además que la distribución de pérdidas de este procedimiento tiene un posible ajuste a cuatro distribuciones, siendo estas; Negativa Binomial, Uniforme, Geométrica y una Poisson. En general una distribución de probabilidad que podría modelar este fenómeno puede ser una Poisson. Esta distribución a pesar de teóricamente representar la tasa de eventos en un período n puede representar este procedimiento como el costo esperado por mes, esto se debe a que se tiene un producto entre una Poisson de entrada y una discreta empírica. Es de esperar que haya gran relación en la frecuencia de entrada y el costo de salida como si fuera casi un producto de una Poisson por una constante, eso se debe a que la gran mayoría de los costos (>70%) tienen un valor único, muy pocos del resto tienen valores lejanos a este costo. Por eso se puede interpretar que la función de pérdidas puede seguir una Poisson, en donde los valores del costo son discretos ya que no hay una infinidad de posibilidades para este.

ID Procedimiento 12211 (879111 - TOMOGRAFIA AXIAL COMPUTADA DE CRANEO SIMPLE):

El OpVar para este procedimiento resultó ser \$17.703.100. De nuevo como se observó en la distribución del procedimiento anterior, al contener una entrada empírica en la severidad con muchos datos pertenecientes a un solo valor, el resultado del histograma muestra un comportamiento similar a la distribución de entrada de la frecuencia, como un producto escalar de la distribución de entrada. Los resultados arrojados por Risk indicaron que la distribución de la función de pérdidas de este proceso se puede modelar como cualquiera de una distribución Geométrica, Negativa Binomial, Uniforme y Poisson.

4. CONCLUSIONES Y CONSIDERACIONES FINALES

4.1. Conclusiones

Mediante el RPA se pudo facilitar la aplicación de la metodología propuesta para tratar grandes cantidades de datos con base en tareas operativas y repetitivas. Esto permite un procesamiento de datos ágil generando caracterizaciones, clasificaciones y clusterizaciones de manera iterativa para diferentes tipos de análisis como la segmentación ABC, las series de tiempo, la caracterización de la morbilidad por grupos, la caracterización de morbilidad general y el OpVar. Este proceso es fácilmente replicable, mantenible y extensible.

Se concluye que existen enfermedades con diagnósticos indicativos que afectan la población en gran medida para Hospitales nivel III en la ciudad de Medellín. Esto se evidenció con los resultados de morbilidad general con los diagnósticos O800 y N189, refiriéndose a partos e insuficiencia renal. En promedio una persona de 100 que entra al hospital se le asigna el diagnóstico O800 mientras tres de cada cien se le asigna el diagnóstico de Insuficiencia renal. Además de esto, estos diagnósticos se ligan a ciertos grupos de etapas de vida específicos como el caso del diagnóstico O800 para el grupo 8 femenino y el N189 para los grupos 12.

Se concluye que a gran parte de la población se les asigna diagnósticos generales los cuales no permiten identificar una enfermedad específica de manera clara. Esto se evidenció para la población general en donde aproximadamente uno de cada diez personas se les asigna el diagnóstico Z139 y más aún no se contienen datos del diagnóstico de egreso. Otro ejemplo es el diagnóstico R688, un diagnóstico que indica otros síntomas y signos generales especificados, en donde cerca de una de cada 20 personas recibe este diagnóstico. Cerca de 20-23% de los diagnósticos son generales, lo cual no permite asignar un apellido directo a la enfermedad, dificultando la caracterización de la morbilidad.

En base a la conclusión anterior y algunos campos como personal que atiende y mortalidad se concluye que para realizar un estudio verdaderamente preciso de la demanda de hospitales los datos requieren mayor especificidad y un mejor diligenciamiento.

Las series de tiempo empleadas permitieron visibilizar el comportamiento de los datos para los procedimientos con id's 9073 y 12211 como para los diagnósticos O800 y Z139. La descomposición

de series de tiempo mediante componentes permitió analizar los patrones de estacionalidad y tendencia de datos permitiendo identificar el comportamiento y tomar decisiones administrativas sobre estos en un futuro.

El OpVar permitió identificar con base en entradas no determinísticas mediante simulación Monte Carlo el capital mínimo requerido para tener solvencia de los procedimientos con IDs 9073 y 12211.

4.2. Recomendaciones

Se recomienda realizar estudios futuros en los cuales además de estudiar los diagnósticos con base en los RIPS, se estudien los procedimientos involucrados para los diagnósticos generales. Esto se debe a que los procedimientos aplicados permiten identificar la sintomatología presentada por el paciente, aunque no permiten asignar un diagnóstico específico, es una forma de aproximarse a entender una porción la demanda de los servicios.

Por otro lado, se recomienda realizar investigaciones respecto a estos diagnósticos generales, como el Z139 para determinar la verdadera morbilidad de estos pacientes y determinar si están sanos o si en realidad no requieren la atención de un Hospital Nivel III.

Se recomienda realizar un análisis estudio de la morbilidad por grupos de vida con el fin de establecer planes de prevención de salud.

Un estudio respecto al procedimiento de diligenciamiento de los RIPS brindaría gran información y podría servir para mejorar la calidad de reporte de estos datos y volver más eficiente este proceso.

Se recomienda desarrollar RPA's con base en el sector salud en donde se busque ahorrar tiempos y agilizar procesos de back office con los cuales se puede buscar mejorar la calidad del servicio y la información brindada.

REFERENCIAS

- Ageron, B., Benzidia, S., & Bourlakis, M. (2018). Healthcare logistics and supply chain – issues and future challenges. *Supply Chain Forum: An International Journal*, 19(1), 1–3. <https://doi.org/10.1080/16258312.2018.1433353>
- Aguirre, S., & Rodriguez, A. (2017). *Automation of a Business Process Using Robotic Process Automation (RPA): A Case Study*. https://doi.org/10.1007/978-3-319-66963-2_7
- Alba, R. D. (1973). A GRAPH-THEORETIC DEFINITION OF A SOCIOMETRIC CLIQUE *. *Journal of Mathematical Sociology* © Gordon and Breach Science Publishers (Vol. 3). Retrieved from <http://aris.ss.uci.edu/~lin/1.pdf>
- Aneiros Pérez, G. (2008). *Series de Tiempo*. Retrieved from <http://eio.usc.es/eipc1/BASE/BASEMASTER/FORMULARIOS-PHP/MATERIALESMATER/Tema1.pdf>
- Ardila, L. (2018). *Notas de Clase*. Medellín.
- Bernal, O., Forero, J. C., Del Pilar Villamil, M., & Pino, R. (2012). Investigación original / Original research Forma de citar Disponibilidad de datos y perfil de morbilidad en Colombia. In *Rev Panam Salud Publica* (Vol. 31). Retrieved from https://www.usfx.bo/nueva/vicerrectorado/citas/TECNOLOGICAS_20/Ingenieria a Sistemas/19.pdf
- Blunck, H., Vican, V., Becker, T., & Windt, K. (2014). Improvement Heuristics for Manufacturing System Design Using Complex Network Figures. *Procedia CIRP*, 17, 50–55. <https://doi.org/10.1016/J.PROCIR.2014.01.063>
- Castro Hoyos, C. E. (2012). *Salud y seguridad social : un breve comparativo de cinco países de América Latina*. Retrieved from <http://library.fes.de/pdf-files/bueros/kolumbien/09229.pdf>
- Castro Zuluaga, C. A., Vélez Gallego, M. C., & Catro Urrego, J. A. (2011). Clasificación ABC Multicriterio: Tipos de Criterios y efectos en la asignación de pesos. *ITECKNE*, 8(2). <https://doi.org/10.15332/iteckne.v8i2.35>
- Choi, T. Y., Dooley, K. J., & Rungtusanatham, M. (2001). *Supply networks and complex adaptive systems: control versus emergence*. *Journal of Operations Management* (Vol. 19). Retrieved from <https://ac-els-cdn-com.recursosbiblioteca.eia.edu.co/S0272696300000681/1-s2.0->

S0272696300000681-main.pdf? tid=de8f44ad-5c84-455a-9406-187b1eafb036&acdnat=1541881817_839e7e5cdbe9010bf8ca104b9639b8e6

Chopra, S., & Meindl, P. (2016). Administración de la cadena de suministro Estrategia, Planeación y Opreación Sexta Edición. Retrieved from www.pearsoned.com/permissions/.

Correa, A. G., Jaime, G., Alzate, G., Luz, C., Escobar, T., Alberto, L., ... Giraldo Velásquez, L. (2012). *Plan de Salud municipal*. Retrieved from: [https://www.medellin.gov.co/irj/go/km/docs/pccdesign/SubportaldeCiudadano_2/PlandeDesarrollo_0_15/InformacinGeneral/SharedContent/Documentos/instrumentos/ps/PLAN DE SALUD MUNICIPAL.pdf](https://www.medellin.gov.co/irj/go/km/docs/pccdesign/SubportaldeCiudadano_2/PlandeDesarrollo_0_15/InformacinGeneral/SharedContent/Documentos/instrumentos/ps/PLAN%20DE%20SALUD%20MUNICIPAL.pdf).

Daniel, W.W. (1999). Biostatistics: A Foundation for Analysis in the Health Sciences. 7th edition. New York: John Wiley & Sons.

Flores, M. (2012). Cuantificación del riesgo operacional mediante modelos de pérdidas agregadas y simulación Monte Carlo. *Analítika*, 5(1), 40–41.

Li, Y., Tao, F., Cheng, Y., Zhang, X., & Nee, A. Y. C. (2017). Complex networks in advanced manufacturing systems. *Journal of Manufacturing Systems*, 43, 409–421. <https://doi.org/10.1016/j.jmsy.2016.12.001>

Li, Y., Cheng, Y., Zhang, X., & Nee, A. Y. C. (2017). Complex networks in advanced manufacturing systems. *Journal of Manufacturing Systems*, 43, 409–421. <https://doi.org/10.1016/J.JMSY.2016.12.001>

Memon, M. S., Lee, Y. H., & Mari, S. I. (2015). Group multi-criteria supplier selection using combined grey systems theory and uncertainty theory. *Expert Systems with Applications*, 42(21), 7951–7959. <https://doi.org/10.1016/J.ESWA.2015.06.018>

Ministerio de Salud. (1994). *Hoja 1 de 1 RESOLUCION NUMERO 5261 DE 1994 (Agosto 5) Por la cual se establece el Manual de Actividades EL MINISTRO DE SALUD*. Retrieved from [https://www.minsalud.gov.co/Normatividad_Nuevo/RESOLUCIÓN 5261 DE 1994.pdf](https://www.minsalud.gov.co/Normatividad_Nuevo/RESOLUCIÓN%205261%20DE%201994.pdf)

MINSALUD. (2015). *LEY ESTATUTARIA No. POR MEDIO DE LA CUAL SE REGULA EL DERECHO FUNDAMENTAL A LA SALUD Y SE DICTAN OTRAS DISPOSICIONES. EL CONGRESO DE COLOMBIA*. Retrieved from [https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/INEC/IGUB/ey-1751-de-2015.pdf](https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/INEC/IGUB/ley-1751-de-2015.pdf)

Ministerio de Salud. (2013). *Documentación reportes RIPS*. Bogotá.

- MINSALUD. (2018). Páginas - Comportamiento del aseguramiento. Retrieved October 3, 2018, from <https://www.minsalud.gov.co/proteccionsocial/Regimensubsidiado/Paginas/coberturas-del-regimen-subsidiado.aspx>
- Ministerio de Salud. (n.d.). Ciclo de Vida. Retrieved May 20, 2019, from <https://www.minsalud.gov.co/proteccionsocial/Paginas/cicloVida.aspx>
- Naciones Unidas. (2018). Salud – Desarrollo Sostenible. Retrieved November 13, 2018, from <https://www.un.org/sustainabledevelopment/es/health/>
- OMS | Estadísticas y salud. (2010). *WHO*. Retrieved from <https://www.who.int/features/qa/73/es/>
- Rentería-Ramos, R., Hurtado Heredía, R., & Piedad Urdinola, B. (2017). Morbimortality of the Victims of Internal Conflict and Poor Population in the Risaralda Province , Colombia.
- Rentería-Ramos, R., & Vitale Alfonso, A. M. (2017). Complex adaptive system in management of hospitality logistis in colombia rafael rentería-ramos 1 alicia maría vitale alfonso 2.
- SERVICIOS DE LA SALUD. (n.d.). Retrieved from http://white.lim.ilo.org/spanish/260ameri/oitreg/activid/proyectos/actrav/proyectos/proyecto_ssos/formacion_distancia/modulo_004/01_capitulo02.pdf
- Volland, J., Fügner, A., Schoenfelder, J., & Brunner, J. O. (2017). Material logistics in hospitals: A literature review. *Omega*, 69, 82–101. <https://doi.org/10.1016/J.OMEGA.2016.08.004>
- WAO. (2016). PAHO. In *ICD 10* (10th ed.). Retrieved from <http://ais.paho.org/classifications/Chapters/CAP21.html>
- Willcocks, L., Lacity, M., & Craig, A. (2017). Robotic Process Automation: Strategic Transformation Lever for Global Business Services? *Journal of Information Technology Teaching Cases*, 7(1), 17–28. <https://doi.org/10.1057/s41266-016-0016-9>
- Wu, S. (2013), "A review on coarse warranty data and analysis", *Reliability Engineering and System*, 114: 1–11. <https://doi:10.1016/j.ress.2012.12.021>

Xia, B. S., & Gong, P. (2015). Review of business intelligence through data analysis. *Benchmarking*, 21(2), 300-311. <https://doi:10.1108/BIJ-08-2012-0050>

ANEXO 1: CÓDIGO SQL

let

```
Source = Excel.CurrentWorkbook(){[Name="Tabla_DatosExternos_13"]}[Content],
```

```
#"Changed Type" = Table.TransformColumnTypes(Source,{{"[Prestación de Servicios de Salud].[$Procedimientos médicos.ID Procedimientos]", Int64.Type}, {"[Prestación de Servicios de Salud].[$Fecha de Atención.Fecha]", type date}, {"[Prestación de Servicios de Salud].[$Personas.Persona Basica ID]", Int64.Type}, {"[Prestación de Servicios de Salud].[$Administradora.Administradora]", type text}, {"[Prestación de Servicios de Salud].[$Diagnóstico Principal.Diagnósticos]", type text}, {"[Prestación de Servicios de Salud].[$Diagnóstico Principal Egreso.Diagnósticos]", type text}, {"[Prestación de Servicios de Salud].[$Ambitos de procedimiento.Ambito del Procedimiento]", type text}, {"[Prestación de Servicios de Salud].[$Finalidad de procedimientos.Finalidad Procedimientos]", type text}, {"[Prestación de Servicios de Salud].[$Finalidad Consulta.Finalidad Consulta]", type text}, {"[Prestación de Servicios de Salud].[$Tipo Usuario.Tipo de Usuario]", type text}, {"[Prestación de Servicios de Salud].[$Personal que atiende.Personal que atiende]", type text}, {"[Prestación de Servicios de Salud].[$Causas externas.Causa Externa]", type text}, {"[Prestación de Servicios de Salud].[$Prestadores Basicos.Prestador Basico ID]", type text}, {"[Prestación de Servicios de Salud].[$Tipo de Atención.Tipo de Atención]", type text}, {"[Prestación de Servicios de Salud].[$Tipo Egreso.Tipo Egreso]", type text}, {"[Prestación de Servicios de Salud].[$Municipio Residencia - RIPS.ID Geografia]", Int64.Type}, {"[Prestación de Servicios de Salud].[$Tipo Diagnóstico Principal.Tipo Diagnóstico Principal]", type text}, {"[Prestación de Servicios de Salud].[$Estado Salida.Condicion Final]", type text}, {"[Prestación de Servicios de Salud].[$Edad de la Atención.Edad]", Int64.Type}, {"[Prestación de Servicios de Salud].[$Número Días de Estancia]", Int64.Type}, {"[Prestación de Servicios de Salud].[$Costo Procedimiento]", Int64.Type}, {"[Prestación de Servicios de Salud].[$Valor Consulta]", Int64.Type}, {"[Prestación de Servicios de Salud].[$Valor Cuota Moderadora]", Int64.Type}, {"[Prestación de Servicios de Salud].[$Neto A Pagar - Consulta]", Int64.Type}, {"[Prestación de Servicios de Salud].[$Valor Cuota Moderadora - Consulta]", Int64.Type}, {"[Prestación de Servicios de Salud].[$Número de Atenciones]", Int64.Type}, {"Genero", type text}}),
```

```
#"Renamed Columns" = Table.RenameColumns("#Changed Type",{{"[Prestación de Servicios de Salud].[$Procedimientos médicos.ID Procedimientos]", "ID Procedimientos"}, {"[Prestación de Servicios de Salud].[$Fecha de Atención.Fecha]", "Semana"}}),
```

```
#"Duplicated Column" = Table.DuplicateColumn("#Renamed Columns", "Semana", "Semana - Copy"),
```

```
#"Renamed Columns6" = Table.RenameColumns("#Duplicated Column",{{"Semana - Copy", "Mes"}}),
```

```
#"Inserted Month" = Table.AddColumn("#Renamed Columns6", "Month", each Date.Month([Mes]), Int64.Type),
```

#"Inserted Year" = Table.AddColumn(#"Inserted Month", "Year", each Date.Year([Mes]), Int64.Type),

#"Removed Columns2" = Table.RemoveColumns(#"Inserted Year",{"Mes"}),

#"Calculated Week of Year" = Table.TransformColumns(#"Removed Columns2",{{"Semana", Date.WeekOfYear, Int64.Type}}),

#"Renamed Columns1" = Table.RenameColumns(#"Calculated Week of Year",{{"[Prestación de Servicios de Salud].\$Personas.Persona Basica ID]", "ID Persona"}, {"[Prestación de Servicios de Salud].\$Administradora.Administradora]", "Administradora"}, {"[Prestación de Servicios de Salud].\$Diagnóstico Principal.Diagnósticos]", "Diagnostico"}}),

#"Filtered Rows" = Table.SelectRows(#"Renamed Columns1", each ([Diagnostico] <> "1 - NO DEFINIDO" and [Diagnostico] <> "Z768 - PERSONA EN CONTACTO CON LOS SERVICIOS DE SALUD EN OTRAS CIRCUNSTANCIAS ESPECIFICADAS")),

#"Renamed Columns2" = Table.RenameColumns(#"Filtered Rows",{{"[Prestación de Servicios de Salud].\$Diagnóstico Principal Egreso.Diagnósticos]", "Diagnostico Egreso"}, {"[Prestación de Servicios de Salud].\$Ambitos de procedimiento.Ambito del Procedimiento]", "Ambito de Procedimiento"}}),

#"Filtered Rows1" = Table.SelectRows(#"Renamed Columns2", each true),

#"Renamed Columns3" = Table.RenameColumns(#"Filtered Rows1",{{"[Prestación de Servicios de Salud].\$Finalidad de procedimientos.Finalidad Procedimientos]", "Finalidad Procedimiento"}, {"[Prestación de Servicios de Salud].\$Finalidad Consulta.Finalidad Consulta]", "Finalidad Consulta"}, {"[Prestación de Servicios de Salud].\$Tipo Usuario.Tipo de Usuario]", "Tipo Usuario"}, {"[Prestación de Servicios de Salud].\$Personal que atiende.Personal que atiende]", "Personal Que Atiende"}, {"[Prestación de Servicios de Salud].\$Causas externas.Causa Externa]", "Causas Externas"}, {"[Prestación de Servicios de Salud].\$Tipo de Atencion.Tipo de Atención]", "Tipo de Atencion"}}),

#"Removed Columns" = Table.RemoveColumns(#"Renamed Columns3",{"[Prestación de Servicios de Salud].\$Tipo Egreso.Tipo Egreso]", "[Prestación de Servicios de Salud].\$Prestadores Basicos.Prestador Basico ID]", "Finalidad Consulta", "[Prestación de Servicios de Salud].\$Tipo Diagnóstico Principal.Tipo Diagnóstico Principal})),

#"Renamed Columns4" = Table.RenameColumns(#"Removed Columns",{{"[Prestación de Servicios de Salud].\$Municipio Residencia - RIPS.ID Geografía]", "Municipio Residencia"}, {"[Prestación de Servicios de Salud].\$Estado Salida.Condicion Final]", "Estado Salida"}, {"[Prestación de Servicios de Salud].\$Edad de la Atención.Edad]", "Edad"}, {"[Prestación de Servicios de Salud].Número Días de Estancia]", "Dias Estancia"}, {"[Prestación de Servicios de Salud].Costo Procedimiento]", "Costo Procedimiento"}, {"[Prestación de Servicios de Salud].Valor Consulta]", "Valor Consulta"}}),

#"Removed Columns1" = Table.RemoveColumns(#"Renamed Columns4",{"[Prestación de Servicios de Salud].Valor Cuota Moderadora]", "[Prestación de Servicios de

Salud].[Neto A Pagar - Consulta]", "[Prestación de Servicios de Salud].[Valor Cuota Moderadora - Consulta]"}},

#"Renamed Columns5" = Table.RenameColumns(#"Removed Columns1",{{"[Prestación de Servicios de Salud].[Número de Atenciones]", "Numero Atenciones"}}),

#"Appended Query" = Table.Combine({#"Renamed Columns5", Raw_2014, Raw_2015, Raw_2016})

in

#"Appended Query"

ANEXO 2: CÓDIGO R

```
# Instalar la librería TSA y Cairo

library(TSA)

library(Cairo)

library(Formula)

library(Hmisc)

Y <- ts(data =
c(28,26,36,30,4,32,20,38,34,26,48,27,13,18,14,30,8,36,9,13,5,7,7,6,10,2,14,20,32,
28,20,23,20,19,24,10,8,29,19,22,14,12,10,11,8,5,8,3), # 2017

      freq = 12,

      start = c(2014,1))

print(Y)

plot (Y, main=expression("Procedimiento 9073"))

minor.tick(nx=12, tick.ratio=0.5 )

#Uso de la función decompose

descompY<-decompose(Y, type = "additive")

plot (descompY)

#Analizar cada componente por separado

#Tendencia

descompY$trend

print(descompY$trend)

plot(descompY$trend, main=expression("Tendencia 9073"))

minor.tick(nx=12, tick.ratio=0.5 )

#Estacionalidad
```

```

descompY$seasonal
plot(descompY$seasonal, main=expression("Estacionalidad 9073"))
minor.tick(nx=12, tick.ratio=0.5 )

#Error
descompY$random
plot(descompY$random, main=expression("Error 9073"))
minor.tick(nx=12, tick.ratio=0.5 )

X <- ts(data=c(1, 0, 0, 0, 0, 0, 0, 0, 4, 0, 0, 0, 757, 858, 921,
              825, 895, 805, 1202, 5656, 11147, 9980, 9720, 3096,
              1048, 1200, 1591, 1323, 2369, 3198, 3629, 2646, 1615,
              1713, 388, 1096, 1710, 1464, 2167, 1961, 2993, 2387, 1530,
              60, 0, 0, 0, 0), freq=12, start =c(2014,1) )

print(X)
plot (X, main=expression("Enfermedad Z139"))
minor.tick(nx=12, tick.ratio=0.5 )

#Uso de la función decompose
descompX <- decompose(X, type = "additive")
plot (descompX)

#Analizar cada componente por separado

#Tendencia
descompX$trend
plot(descompX$trend, main=expression("Tendencia Z139"))
minor.tick(nx=12, tick.ratio=0.5 )

#Estacionalidad

```

```

descompX$seasonal
plot(descompX$seasonal, main=expression("Estacionalidad Z139"))
minor.tick(nx=12, tick.ratio=0.5 )
#Error
descompX$random
plot(descompX$random, main=expression("Error Z139"))
Z <- ts(data=c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
              8, 8, 8, 9, 2, 13, 8, 40, 100, 96, 79,
              19, 3, 0, 7, 6, 15, 8, 18, 13, 6, 4, 0,
              11, 6, 5, 4, 9, 26, 19, 7, 0, 0, 0, 0, 0), freq=12, start=c(2014,1) )
print(Z)
plot (Z, main=expression("Procedimiento 12211"))
minor.tick(nx=12, tick.ratio=0.5 )
#Uso de la función decompose
descompZ <- decompose(Z, type = "additive")
plot (descompZ)
#Tendencia
descompZ$trend
plot(descompZ$trend, main=expression("Tendencia 12211"))
minor.tick(nx=12, tick.ratio=0.5 )
#Estacionalidad
descompZ$seasonal
plot(descompZ$seasonal, main=expression("Estacionalidad 12211"))
minor.tick(nx=12, tick.ratio=0.5 )

```

```

#Error

descompZ$random

plot(descompZ$random, main=expression("Error 12211"))

a <- ts(data=c(84, 83, 134, 99, 18, 109, 104, 148, 114, 140, 211, 105, 131, 88, 98,
119, 90, 185, 88, 121, 117, 88, 87, 104, 156, 97, 133, 149, 154, 123, 102, 139,
145, 104, 82, 53, 44, 173, 144, 136, 84, 90, 51, 53, 71, 42, 56, 39
), freq=12, start =c(2014,1) )

print(a)

plot (a, main=expression("Enfermedad O800"))

minor.tick(nx=12, tick.ratio=0.5 )

#Uso de la función decompose

descompa <- decompose(a, type = "additive")

plot (descompa)

#Tendencia

descompa$trend

plot(descompa$trend, main=expression("Tendencia O800"))

minor.tick(nx=12, tick.ratio=0.5 )

#Estacionalidad

descompa$seasonal

plot(descompa$seasonal, main=expression("Estacionalidad O800"))

minor.tick(nx=12, tick.ratio=0.5 )

#Error

descompa$random

plot(descompa$random, main=expression("Error O800"))

```

ANEXO 3: Datos Severidad

ID Procedimiento	Mes	Costo Procedimiento	ID Procedimiento	Mes	Costo Procedimiento
		\$			\$
9073	1/01/2017	542.385	12211	5/2017	411.700
		\$			\$
9073	1/01/2017	542.385	12211	5/2017	411.700
		\$			\$
9073	1/01/2017	542.385	12211	5/2017	411.700
		\$			\$
9073	1/01/2017	542.385	12211	5/2017	411.700
		\$			\$
9073	1/01/2017	638.100	12211	5/2017	411.700
		\$			\$
9073	1/01/2017	542.385	12211	5/2017	411.700
		\$			\$
9073	1/01/2017	542.385	12211	5/2017	411.700
		\$			\$
9073	1/01/2017	542.385	12211	4/2017	411.700
		\$			\$
9073	1/02/2017	638.100	12211	1/2017	411.700
		\$			\$
9073	1/02/2017	638.100	12211	3/2017	411.700
		\$			\$
9073	1/02/2017	638.100	12211	3/2017	411.700
		\$			\$
9073	1/02/2017	638.100	12211	3/2017	411.700
		\$			\$
9073	1/02/2017	542.385	12211	5/2017	411.700
		\$			\$
9073	1/02/2017	542.385	12211	5/2017	411.700
		\$			\$
9073	1/02/2017	542.385	12211	6/2017	411.700
		\$			\$
9073	1/02/2017	542.385	12211	5/2017	411.700
		\$			\$
9073	1/02/2017	542.385	12211	5/2017	411.700
		\$			\$
9073	1/02/2017	542.385	12211	7/2017	411.700
		\$			\$
9073	1/02/2017	542.385	12211	6/2017	411.700
		\$			\$
9073	1/02/2017	542.385	12211	6/2017	411.700

		\$			\$
9073	1/02/2017	542.385		12211	2/2017
		\$			\$
9073	1/02/2017	542.385		12211	4/2017
		\$			\$
9073	1/02/2017	542.385		12211	6/2017
		\$			\$
9073	1/02/2017	542.385		12211	4/2017
		\$			\$
9073	1/02/2017	542.385		12211	4/2017
		\$			\$
9073	1/02/2017	542.385		12211	5/2017
		\$			\$
9073	1/02/2017	542.385		12211	6/2017
		\$			\$
9073	1/02/2017	638.100		12211	4/2017
		\$			\$
9073	1/02/2017	542.385		12211	4/2017
		\$			\$
9073	1/02/2017	542.385		12211	5/2017
		\$			\$
9073	1/02/2017	542.385		12211	5/2017
		\$			\$
9073	1/02/2017	542.385		12211	6/2017
		\$			\$
9073	1/02/2017	542.385		12211	5/2017
		\$			\$
9073	1/02/2017	542.385		12211	6/2017
		\$			\$
9073	1/02/2017	542.385		12211	6/2017
		\$			\$
9073	1/02/2017	542.385		12211	7/2017
		\$			\$
9073	1/03/2017	593.433		12211	6/2017
		\$			\$
9073	1/03/2017	638.100		12211	6/2017
		\$			\$
9073	1/03/2017	542.385		12211	2/2017
		\$			\$
9073	1/03/2017	542.385		12211	5/2017
		\$			\$
9073	1/03/2017	542.385		12211	5/2017

		\$			\$	
9073	1/03/2017	542.385		12211	5/2017	411.700
		\$				\$
9073	1/03/2017	542.385		12211	5/2017	411.700
		\$				\$
9073	1/03/2017	542.385		12211	5/2017	411.700
		\$				\$
9073	1/03/2017	542.385		12211	5/2017	411.700
		\$				\$
9073	1/03/2017	542.385		12211	5/2017	411.700
		\$				\$
9073	1/03/2017	542.385		12211	1/2017	411.700
		\$				\$
9073	1/03/2017	542.385		12211	6/2017	411.700
		\$				\$
9073	1/03/2017	542.385		12211	4/2017	411.700
		\$				\$
9073	1/03/2017	542.385		12211	1/2017	411.700
		\$				\$
9073	1/03/2017	542.385		12211	3/2017	411.700
		\$				\$
9073	1/03/2017	542.385		12211	7/2017	411.700
		\$				\$
9073	1/03/2017	542.385		12211	5/2017	411.700
		\$				\$
9073	1/03/2017	542.385		12211	7/2017	411.700
		\$				\$
9073	1/03/2017	1.084.770		12211	6/2017	411.700
		\$				\$
9073	1/04/2017	542.385		12211	4/2017	411.700
		\$				\$
9073	1/04/2017	542.385		12211	6/2017	411.700
		\$				\$
9073	1/04/2017	542.385		12211	2/2017	411.700
		\$				\$
9073	1/04/2017	542.385		12211	4/2017	411.700
		\$				\$
9073	1/04/2017	542.385		12211	1/2017	411.700
		\$				\$
9073	1/04/2017	542.385		12211	6/2017	411.700
		\$				\$
9073	1/04/2017	542.385		12211	6/2017	411.700
		\$				\$
9073	1/04/2017	542.385		12211	5/2017	411.700

		\$			\$
9073	1/04/2017	542.385		12211	5/2017
		\$			411.700
9073	1/04/2017	542.385		12211	2/2017
		\$			411.700
9073	1/04/2017	542.385		12211	5/2017
		\$			411.700
9073	1/04/2017	542.385		12211	6/2017
		\$			411.700
9073	1/04/2017	542.385		12211	7/2017
		\$			411.700
9073	1/04/2017	542.385		12211	6/2017
		\$			411.700
9073	1/04/2017	542.385		12211	7/2017
		\$			411.700
9073	1/04/2017	542.385		12211	1/2017
		\$			411.700
9073	1/04/2017	542.385		12211	1/2017
		\$			411.700
9073	1/04/2017	542.385		12211	6/2017
		\$			411.700
9073	1/04/2017	542.385		12211	7/2017
		\$			411.700
9073	1/04/2017	542.385		12211	2/2017
		\$			411.700
9073	1/04/2017	542.385			
		\$			
9073	1/04/2017	638.100			
		\$			
9073	1/05/2017	593.433			
		\$			
9073	1/05/2017	542.385			
		\$			
9073	1/05/2017	542.385			
		\$			
9073	1/05/2017	542.385			
		\$			
9073	1/05/2017	542.385			
		\$			
9073	1/05/2017	542.385			
		\$			
9073	1/05/2017	542.385			
		\$			
9073	1/05/2017	542.385			

		\$
9073	1/05/2017	542.385
		\$
9073	1/05/2017	542.385
		\$
9073	1/05/2017	638.100
		\$
9073	1/05/2017	638.100
		\$
9073	1/05/2017	542.385
		\$
9073	1/05/2017	542.385
		\$
9073	1/06/2017	542.385
		\$
9073	1/06/2017	542.385
		\$
9073	1/06/2017	542.385
		\$
9073	1/06/2017	542.385
		\$
9073	1/06/2017	542.385
		\$
9073	1/06/2017	542.385
		\$
9073	1/06/2017	542.385
		\$
9073	1/06/2017	638.100
		\$
9073	1/06/2017	542.385
		\$
9073	1/06/2017	542.385
		\$
9073	1/07/2017	542.385
		\$
9073	1/07/2017	542.385
		\$
9073	1/07/2017	638.100
		\$
9073	1/07/2017	542.385

		\$
9073	1/07/2017	542.385
		\$
9073	1/07/2017	638.100
		\$
9073	1/07/2017	1.276.200
		\$
9073	1/07/2017	542.385
		\$
9073	1/07/2017	542.385
		\$
9073	1/07/2017	542.385
		\$
9073	1/08/2017	542.385
		\$
9073	1/08/2017	542.385
		\$
9073	1/08/2017	542.385
		\$
9073	1/08/2017	542.385
		\$
9073	1/08/2017	542.385
		\$
9073	1/08/2017	542.385
		\$
9073	1/08/2017	1.084.770
		\$
9073	1/08/2017	1.084.770
		\$
9073	1/08/2017	1.084.770
		\$
9073	1/08/2017	1.084.770
		\$
9073	1/08/2017	1.084.770
		\$
9073	1/09/2017	542.385
		\$
9073	1/09/2017	1.084.770
		\$
9073	1/09/2017	1.084.770
		\$
9073	1/09/2017	1.084.770
		\$
9073	1/09/2017	542.385

		\$
9073	1/09/2017	542.385
		\$
9073	1/09/2017	638.100
		\$
9073	1/09/2017	542.385
		\$
9073	1/10/2017	434.358
		\$
9073	1/10/2017	542.385
		\$
9073	1/10/2017	587.052
		\$
9073	1/10/2017	542.385
		\$
9073	1/10/2017	542.385
		\$
9073	1/11/2017	542.385
		\$
9073	1/11/2017	542.385
		\$
9073	1/11/2017	542.385
		\$
9073	1/11/2017	542.385
		\$
9073	1/11/2017	542.385
		\$
9073	1/11/2017	638.100
		\$
9073	1/11/2017	542.385
		\$
9073	1/11/2017	542.385
		\$
9073	1/12/2017	542.385
		\$
9073	1/12/2017	542.385
		\$
9073	1/12/2017	434.358

